

***The genetics of drug resistance in malaria
- identification of genes conferring
chloroquine and artemisinin resistance in
rodent malaria parasite Plasmodium
chabaudi.***

Katarzyna Kinga Modrzynska

Thesis submitted for a degree of Doctor of Philosophy
University of Edinburgh
June 2010

I hereby declare that, except the contributions specified in the following sections, the work presented here was planned, performed and written by myself only and has not been submitted for any other degree or professional qualification.

Katarzyna Kinga Modrzynska

June 2010

Acknowledgements

This work would not be possible without a help and support of many people, to whom I would like to thank here.

The main credit goes to dr Paul Hunt who for unexplained reasons took me as his PhD student and coped with me for four years, always showing understanding and support far exceeding the supervisor duties. Prof. Richard Carter, my second supervisor, provided the mentorship, inspiration and vast knowledge of rodent malaria system throughout the whole project. Mr Richard Fawcett trained me in mouse work and provided excellent technical assistance during all mouse experiments – many happy and silly moments were spend together in the mousehouse. The arrival of dr Axel Martinelli from Portugal provided a missing link between informatics and biology and made the Solexa sequencing happen. Dr Alison Creasy provided help with pyrosequencing, molecular biology assistance and a moral support. Dr Laurence Lowe and Dario Beraldi translated the complexity of biological data into a simple language of statistics. Finally, all my officemates: Elaine O'mahoney, Kathlyn Degnan, Sittiporn Pattaradilokrat, Sophia Borges, Carol Hunja, Naomi Gadsby, Louise Rodrigues, Laurence Hall and Laura Jeacok; shared with me the ups and downs of the research live and provided a cheerful and stimulating scientific environment.

I am also very grateful to my parents – Lidia and Andrzej Modrzynscy and parents-in-law – Mariarita Fabri and Enzo Beraldi whose help and support allowed me to cope with challenges of science and motherhood.

Most importantly however I would like to thank my family: my husband Dario who supported me throughout all periods of hysteria, euphoria, depression and excitement that I was passing through during my thesis work and to my daughter Izabella who coped amazingly well with an extensive parental neglect I was subjecting her to during the final stages of completion of this thesis.

Thank you all!

Statement of contributions

The following major contributions to this work have to be acknowledged.

- the general approach to gene identification explored in this thesis was devised by dr Paul Hunt
- the Linkage Group Selection used in Chapters 3 and 4 was developed by prof. Richard Carter and his collaborators.
- the design of pyrosequencing markers used in chapter 3 and 4 was a collaborative project between Paul Hunts' and Richard Carter's groups to which many people contributed (in particular Ms Kathryn Degnan, Dr Sandra Cheesman and dr Alison Creasey have to be credited)
- the post-Solexa genome re-assembly and initial data analysis, described in Chapter 5, was performed by Urmi Trivedi, Sujai Kumar and Axel Martinelli under supervision of dr Paul Hunt
- the statistical analysis of data in Chapter 4 was performed by Dr Laurence Loewe and dr Dario Beraldi
- the statistical analysis of data in chapter 6 was performed by dr Dario Beraldi
- the validation of mutations described in chapter 5 in AS-15MF strain was done by Dr Sofia Borges

Table of contents

ACKNOWLEDGEMENTS	III
STATEMENT OF CONTRIBUTIONS	IV
TABLE OF CONTENTS	V
LIST OF FIGURES	X
LIST OF TABLES	XI
LIST OF ABBREVIATIONS	XII
ABSTRACT	1
1. INTRODUCTION	2
1.1. Malaria	2
1.1.1. The burden of malaria	2
1.1.2. The parasite, vector and life cycle	3
1.1.3. Clinical picture of disease	5
1.1.4. Disease control strategies	6
1.2. Malaria chemotherapy	8
1.2.1. Antifolates	8
1.2.2. Aminoquinolines and chloroquine (CQ)	9
1.2.3. 4-methanolquinolines	12
1.2.4. Artemisinins	12
1.2.5. Atovaquone	13
1.2.6. Antibiotics	13
1.3. Drug resistance in malaria	14
1.3.1. Assessing resistance	14
1.3.2. Evolution of resistance	16
1.3.3. Fitness cost of drug resistance	18
1.4. Chloroquine (CQ) resistance	22
1.4.1. Appearance and spread of CQ resistance	22
1.4.2. Genetics of CQ resistance	22
1.4.3. Molecular mechanisms of CQ resistance	28
1.5. Artemisinin resistance	31
1.5.1. The resistance to artemisinins – an incoming threat	31
1.5.2. The emerging artesunate resistance in central Asia	31
1.5.3. Genetics of ART resistance	32
1.6. Identification of genetic markers of resistance	34
1.6.1. Why are we interested in resistance markers?	34
1.6.2. Methods of marker identification.	34

1.7.	The <i>P.chabaudi</i> model of CQ and ART resistance	37
1.7.1.	Why use the rodent model?	37
1.7.2.	The lineage of resistant <i>P.chabaudi</i>	38
1.7.3.	Past investigations of the genetics of drug resistance in <i>P.chabaudi</i> lineage	40
1.8.	New gene identification strategy – a combination of Linkage Group Selection and next generation sequencing on isogenic drug resistant lines of parasite	43
1.8.1.	Linkage Group Selection – a fast method to map resistance loci	43
1.8.2.	Next generation sequencing – a method of mutation discovery	47
1.9.	Aims and content of this thesis	49
2.	THE PRELIMINARY EXPERIMENTS – ESTABLISHING THE SELECTION CONDITIONS FOR LGS EXPERIMENTS	50
2.1.	Introduction	50
2.1.1.	The importance of the selection in LGS	50
2.1.2.	The requirements of different drugging regimes	51
2.2.	Experimental design	52
2.2.1.	AS-30CQ+AJ	52
2.2.2.	AS-3CQ+AJ	53
2.3.	Materials and methods	57
2.3.1.	Parasite and mice strains	57
2.3.2.	Setting up the experimental infections	57
2.3.3.	Drugging	58
2.3.4.	Monitoring the parasitaemia	58
2.3.5.	Proportional sequencing of <i>pcdhps</i> gene	59
2.4.	Results	62
2.4.1.	AS-30CQ+AJ competition experiment	62
2.4.2.	The AS-3CQ +AJ competition experiment	67
2.5.	Discussion	73
2.5.1.	The confirmation of the phenotypes of studied strains	73
2.5.2.	The drugging regime to select the highly resistant recombinants from low resistant and sensitive ones	74
2.5.3.	The drugging regime to select all resistant recombinants from sensitive ones	74
2.5.4.	The fast-growing sensitive parent: implications for the cross	75
2.5.5.	Possible fitness cost of chloroquine resistance in <i>P.chabaudi</i>	76
3.	GENERATION AND LGS ANALYSIS OF AS-30CQ X AJ CROSS	77
3.1.	Introduction	77
3.1.1.	The resistant and sensitive strain	77
3.1.2.	LGS drug selection conditions	78
3.1.3.	Pyrosequencing	78
3.2.	The experimental design	80
3.2.1.	The generation of the cross	80
3.2.2.	The amplification of the cross progeny	80
3.2.3.	Selection	81
3.2.4.	Genetic analysis	81

3.3. Materials and methods	83
3.3.1. Parasites, mice and mosquitoes:	83
3.3.2. Mouse infections	83
3.3.3. Mosquito feeds	83
3.3.4. Dissection for oocysts	83
3.3.5. The recovery of sporozoites from salivary glands	84
3.3.6. The proportional sequencing of <i>dhps</i> gene	84
3.3.7. Large scale DNA isolation	84
3.3.8. Estimation of frequencies of genetic markers using the pyrosequencing	84
3.4. Results	88
3.4.1. Generation of the AS-30CQxAJ cross	88
3.4.2. Selection of the cross	89
3.4.3. The quantitative genetic analysis of the initial cross	91
3.5. Discussion	95
3.5.1. The AS-30CQxAJ cross	95
3.5.2. The untreated population – the selection for multiple loci for virulence?	95
3.5.3. The treated population – many loci under selection or sampling error?	96
3.5.4. The loss of recombinants?	97
4. GENERATION AND GENETIC ANALYSIS OF AS-30CQ X AJ BACKCROSS	99
4.1. Introduction	99
4.1.1. Rationale for backcross	99
4.1.2. Modifications of LGS protocol- aiming for a higher number of recombinants	100
4.1.3. Modifications of selection procedure – the use of low, intermediate and high drug concentration	101
4.2. Experimental design	102
4.2.1. Pre-selection of the cross and genetic analysis of obtained material	102
4.2.2. Generation of the backcross	102
4.2.3. Selection 1	103
4.2.4. Selection 2	104
4.2.5. The genetic analysis of backcross progeny	104
4.3. Materials and Methods	107
4.3.1. EPO treatment	107
4.3.2. The recovery of the cross progeny from liquid nitrogen	107
4.3.3. Artemisinin treatment	107
4.3.4. Statistical analysis	107
4.4. Results	109
4.4.1. Reselection of cross progeny	109
4.4.2. The generation of the backcross	111
4.4.3. The selection of the backcross	112
4.4.4. The genetic analysis of the backcross – selection I	114
4.4.5. The genetic analysis of the backcross – selection 2	120
4.5. Discussion	131
4.5.1. The improvement in the quality of data	131
4.5.2. Three levels of resistance to chloroquine conferred by three loci?	133
4.5.3. Artemisinin resistance – confirmation of existence of the locus on chromosome 2	135
4.5.4. The possible locus of ART resistance on chr 11	136

5. THE SPECIFIC MUTATIONS CONNECTED WITH CQ AND ART RESISTANCE	138
5.1. Introduction	138
5.2. The experimental design	141
5.3. Materials and Methods	143
5.3.1. The preparation of DNA	143
5.3.2. Illumina-Solexa genome re-sequencing:	143
5.3.3. The bioinformatic analysis of Solexa genome re-sequencing data	143
5.3.4. Dideoxy sequencing	146
5.3.5. Proportional sequencing	146
5.3.6. Protein characterisation	148
5.4. Results	149
5.4.1. The genome re-sequencing and Solexa data analysis	149
5.4.2. The validation of Solexa data	156
5.4.3. The summary of genetic changes between AS-sens and AS-30CQ	158
5.4.4. The candidate mutations associated with CQ/ART resistance and their origin in the lineage	159
5.4.5. Proportional sequencing of predicted mutations in the selected backcross progeny	169
5.4.6. The predicted determinants of CQ and ART resistance	173
5.5. Discussion	187
5.5.1. The mutations within the resistant AS lineage	187
5.5.2. The structure and acquisition of CQ and ART resistance	188
5.5.3. The possible mechanism of CQ/ART resistance in <i>P.chabaudi</i>	190
6. THE FITNESS COST OF HIGH CQ RESISTANCE.	193
6.1. Introduction	193
6.2. Experimental design	194
6.3. Materials and methods	196
6.3.1. Experimental infections	196
6.3.2. Proportional sequencing of ubp1 gene	196
6.3.3. Statistical analysis	197
6.4. Results	198
6.4.1. Single clone infections	198
6.4.2. Mixed infections	200
6.5. Discussion	205
6.5.1. Do mutations underlying chloroquine resistance incur fitness costs?	205
6.5.2. What genetic changes account for the fitness costs?	206
6.5.3. Additional observations	207
7. GENERAL DISCUSSION AND FINAL REMARKS	209
7.1. Summary of results	209
7.2. Combination of LGS and Solexa sequencing for the identification of genes involved in drug resistance	210

7.3.	Relevance for human malaria	211
7.3.1.	The new gene candidates involved in CQ/ART resistance.	211
7.3.2.	The insights into the mechanism of CQ and ART resistance	212
7.3.3.	The connection between chloroquine and artemisinin resistance	213
7.3.4.	A platform for identification of genes connected to drug resistance	214
7.4.	Feature directions	215
7.4.1.	The use of Solexa SNPs quantification to improve the existing gene finding strategy.	215
7.4.2.	Further investigation of gene candidates.	215
7.4.3.	Exploration of fitness cost and its consequences	216
APPENDIX A – BUFFERS AND SOLUTIONS		217
APPENDIX B – GENERAL PCR PROTOCOL.		218
APPENDIX C – THE LIST OF PYROSEQUENCING ASSAYS USED IN LGS ANALYSIS		219
APPENDIX D – THE ALIGNMENT OF THE THREE CONSERVED GENES CONNECTED TO CQ/ART RESISTANCE.		225
APPENDIX E – TALKS, POSTERS AND PUBLICATIONS ARISING FROM THIS THESIS		240
BIBLIOGRAPHY		242

List of Figures

Figure 1.1 The distribution of malaria worldwide.....	2
Figure 1.2 The life cycle of <i>P.falciparum</i>	4
Figure 1.3 The chemical structure of CQ and the putative mechanism of its action.	11
Figure 1.4 Resistance observed <i>in vitro</i>	15
Figure 1.5 Fitness cost of drug resistance.	19
Figure 1.6 <i>Pfcr</i> – the main determinant of CQ resistance in <i>P.falciparum</i>	24
Figure 1.7 <i>PFMDR1</i> – a multidrug resistance protein.....	25
Figure 1.8 The possible role of PFCRT in CQ resistance	30
Figure 1.9 The resistant AS lineage of <i>P.chabaudi</i>	39
Figure 1.10 The design of a Linkage Group Selection experiment	44
Figure 1.11 The possible results of genetic analysis	46
Figure 2.1 The design of AS-30CQ+ AJ experiment	55
Figure 2.2. The design of AS-3CQ+ AJ experiment	56
Figure 2.3 The proportional sequencing.....	61
Figure 2.4 The controls for AS-30CQ + AJ infections.....	65
Figure 2.5 The mixed infections with AS-30CQ and AJ strains in untreated animals.	65
Figure 2.6 The mixed infections with AS-30CQ and AJ strains	66
Figure 2.7 The controls for AS-3CQ + AJ infections.....	70
Figure 2.8 The mixed infections with AS-3-CQ and AJ strains treated with 0 mg of CQ kg ⁻¹ day ⁻¹ ... 71	71
Figure 2.9 The mixed infections with AS-3-CQ and AJ strains treated with 1.5 mg of CQ kg ⁻¹ day ⁻¹ . 71	71
Figure 2.10 The mixed infections with AS-3-CQ and AJ strains treated with 3 mg of CQ kg ⁻¹ day ⁻¹ . 72	72
Figure 2.11 The mixed infections with AS-3-CQ and AJ strains treated with 5 mg of CQ kg ⁻¹ day ⁻¹ . 72	72
Figure 3.1 The use of pyrosequencing to calculate the proportion of SNP based genetic markers in the population.	79
Figure 3.2 The schematic diagram of AS-30CQxAJ LGS experiment	82
Figure 3.3 The parasitaemia of AS-30CQxAJ cross	90
Figure 3.4 The results of genome scan of AS-30CQxAJ cross.	94
Figure 4.1 The improvement of data quality by the backcross.....	100
Figure 4.2 Predicted patterns of appearance of selection valleys in an experimental population.	101
Figure 4.3 The backcross and selection 1 of its progeny.....	105
Figure 4.4 Selection 2 of the progeny of the backcross.....	106
Figure 4.5 The results of genome-scans performed on preselected populations.	110
Figure 4.6 The growth of AS30CQ x AJ backcross during the selection I	113
Figure 4.7 The growth of AS-30CQ x AJ backcross during selection II.....	113
Figure 4.8 Whole genome scans (selection 1)	117
Figure 4.9 Selection 1 transformed data.....	119
Figure 4.10 Selection 2 of AS-30CQ x AJ backcross - untreated	124
Figure 4.11 Selection 2 of AS-30CQ x AJ backcross – treated with 1.5 mg of CQ kg ⁻¹ day ⁻¹	125
Figure 4.12 Selection 2 of AS-30CQ x AJ backcross – treated with 3 mg of CQ kg ⁻¹ day ⁻¹	126
Figure 4.13 Selection 2 of AS-30CQ x AJ backcross – treated with 10 mg of CQ kg ⁻¹ day ⁻¹	127
Figure 4.14 Selection 2 of AS-30CQ x AJ backcross – treated with 20 mg of CQ kg ⁻¹ day ⁻¹	128
Figure 4.15 Selection 2 of AS-30CQ x AJ backcross – treated with 100 mg of ART kg ⁻¹ day ⁻¹	129
Figure 4.16 Selection 2 of AS-30CQ x AJ backcross – treated with 200 mg of ART kg ⁻¹ day ⁻¹	130
Figure 4.17 A model of how mutations may confer CQ resistance in AS lineage.	134
Figure 5.1 Point mutation resulting in A173E substitution in PCHAS_112780 (<i>aat1</i>) gene in various strains of AS lineage.....	160
Figure 5.2 Point mutation resulting in T707N substitution in PCHAS_030200 gene in various strains of AS lineage	161
Figure 5.3 Point mutation resulting in T719N substitution in PCHAS_031370 (<i>tm12</i>) gene in various strains of AS lineage.....	162
Figure 5.4 The 3bp deletion in PCHAS_031370 (<i>tm12</i>) gene in various strains of AS lineage.....	163
Figure 5.5 Point mutation resulting in V2729F substitution in PCHAS_020720 (<i>ubp1</i>) gene in various strains of AS lineage.....	164

Figure 5.6 Results of previous sequencing of the point mutation resulting in V2697F substitution in <i>ubp1</i> gene in various strains of AS lineage.....	165
Figure 5.7 The proportions of the mutations identified in <i>aat1</i> , <i>tm12</i> and <i>ubp1</i> genes in backcross progeny selected with 3, 10 and 20 mg of CQ.....	171
Figure 5.8 The proportions of the mutations identified in <i>aat1</i> , <i>tm12</i> and <i>ubp1</i> genes in backcross progeny selected with 100 and 200 mg ART.....	172
Figure 5.9 The predicted structure of <i>AAT1</i> protein and the expression profile of its <i>P.falciparum</i> orthologue.....	174
Figure 5.10 The fragment of alignment of homologues of <i>AAT1</i> protein between six Plasmodium species.....	175
Figure 5.11 The predicted structure of <i>TM12</i> protein and the expression profile of its <i>P.falciparum</i> homologue.....	177
Figure 5.12 The alignment of two fragments of <i>TM12</i> protein between six Plasmodium species.	179
Figure 5.13 The impact of I102del in <i>TM12</i> on protein structure	180
Figure 5.14 The predicted structure of the fragment of <i>UBP1</i> protein and the expression profile of its <i>P.falciparum</i> homologue.	184
Figure 5.15 The fragment of the alignment of <i>ubp1</i> protein between six Plasmodium species	186
Figure 5.16 The predicted impact of the V2728F and V2697F mutations on the structure of the catalytic domain of <i>UBP1</i>	186
Figure 5.17 The possible order of occurrence of mutations on chr 3 and chr 2 in the AS lineage	189
Figure 6.2 The single strain infections	199
Figure 6.3 AS-30CQ +AS-3CQ mixed infection	202
Figure 6.4 The variation with animals treated with 3mg of CQ kg ⁻¹ day ⁻¹	202
Figure 6.5 Differences between single and mixed infections.....	204

List of Tables

Table 3.1 The results of mosquito dissections for AS-30CQxAJ cross.....	88
Table 3.2 The results of proportional sequencing of recovered progeny from AS-30CQxAJ cross. ...	89
Table 4.1 Oocyst analysis during backcross generation.....	111
Table 5.1 The initial list of gene candidates on chromosome 11	140
Table 5.2 Primers used for validation of sequencing mutations.....	147
Table 5.3 The summary of Solexa sequencing metrics of AS-sens and AS-30CQ strain.	150
Table 5.4 The summary of the coverage of AS-30CQ by chromosome in SSAHA2.....	150
Table 5.5 Number of point mutations identified on each step of analysis with MAQ software.....	152
Table 5.6 The point mutations identified with MAQ software and the results of their visual confirmation.....	153
Table 5.7 The point mutations identified with SSAHA2 software.....	153
Table 5.8 The list of indels identified by bioinformatic analysis.	155
Table 5.9 The validation of the analysis of Solexa results.	157
Table 5.10 Predicted genetic changes between the AS-sens and AS-30CQ.....	158
Table 5.11 The results of genotyping of CQ resistant lineage for mutations found on chr 11, chr 3 and chr 2.....	168
Table 5.12 The percentages of wt AJ allele within backcross progeny.....	169
Table 5.13 Polymorphisms in PFF1430c gene in <i>P.falciparum</i> population.....	175
Table 5.14 Polymorphisms in PFB0675w gene in <i>P.falciparum</i> population	178
Table 5.15 Polymorphisms in PFA0220w gene in <i>P.falciparum</i> population	185

List of abbreviations

General

Aat1 – amino acid transporter 1, a code for PCHAS_112780 gene

ACT – artemisinin combination therapy

ART – artemisinin

bp – base pairs

CQ – chloroquine

dhfr - Dihydrofolate reductase,

dhps - dihydropteroate synthetase

DMSO - Dimethyl sulfoxide

DNA - deoxyribonucleic acid

EPO - erythropoietin

FPXI- ferriprotoporphyrin IX

Kb - kilobases

LGS- Linkage Group Selection

MF- mefloquine

mg kg⁻¹ day⁻¹ – milligrams of drug per kilogram mouse body weight per day

PABA- para-aminobenzoic acid

pc – Plasmodium chabaudi

pf – Plasmodium falciparum

Pfcr1 – P.falciparum chloroquine resistance transporter

Pfmdr1 - P.falciparum multidrug resistance gene

PCR – polymerase chain reaction

p.i. – post inoculation

PYR – pyrimethamine

RBC - red blood cell

Tm12 – transmembrane protein with 12 TM domains, a code for PCHAS_031370 gene

TMD – transmembrane domain

Ubp1- deubiquitinating protease 1

WHO – World Health Organisation

WTSI – Wellcome Trust Sanger Institute

DNA bases

A- adenine

C- cytosine

G- guanine

T- thymine

Amino acids:

A- Alanine

E - Glutamic acid

F - Phenylalanine

I - Isoleucine

N - Asparagine

S - Serine

T - Threonine

V - Valine

Abstract

Resistance to antimalarial drugs continues to be a major obstacle in controlling and eradicating malaria. The identification of genetic markers of resistance is vital for disease management but they can be difficult to predict before resistance arises in the field. This thesis describes an alternative approach to gene identification, combining an *in vivo* experimental evolution model, Linkage Group Selection (LGS) and Solexa genome re-sequencing. Here this model was used to resolve the genetic basis of chloroquine and artemisinin resistance in the rodent malaria parasite *Plasmodium chabaudi*.

AS-30CQ is a parasite with high resistance to chloroquine and resistance to artemisinin. It was crossed with the genetically different drug-sensitive strain AJ. The resulting progeny were selected with drugs and backcrossed to the sensitive parent. Both crosses were treated with increasing concentrations of chloroquine and artemisinin. The frequency of markers from the sensitive parasite were analysed in order to characterize the signatures of drug selection. Three loci involved progressively in chloroquine resistance were identified on chromosomes 11, 3 and 2. One main locus on chromosome 2 was identified with artemisinin selection.

The Solexa platform was used to re-sequence the genomes of both AS-30CQ and its sensitive progenitor, AS-sens. The differences between the two genomes were integrated with the LGS data to identify: 1) a strong candidate for the main CQ-resistance determinant - a putative amino acid transporter on chromosome 11 (*aat1*) 2) two candidates for high level chloroquine resistance on chromosome 3. and 3) a mutation in *ubp1* gene on chromosome 2 that is likely to contribute to the highest level of chloroquine resistance and be main determinant of the artemisinin resistance phenotype.

In addition the last section of this thesis describes two otherwise isogenic clones showing low- and high levels of chloroquine resistance were grown competitively to evaluate the effect of these mutations on parasite fitness. The highly resistant strain demonstrated a loss of fitness in relation to its more sensitive progenitor and was outcompeted in untreated and low-treated infections.

1. Introduction

1.1. *Malaria*

1.1.1. The burden of malaria

Malaria is one of the oldest diseases known to mankind. First reports about the so-called “swamp disease” come from ancient China 2nd century BC, although malaria would have been around for much longer and probably is as old as human species (Hughes and Verra, 2010). Today, it still remains one of the most deadly tropical diseases, taking each year ~1 million lives and causing ~250 million infections (WHO, 2008). Currently ~50% of the world’s population lives in areas of potential malaria transmission. It is widespread throughout of the tropics (Fig 1.1): however most infections, morbidity and mortality occur in sub-Saharan Africa. The most susceptible persons are young children and travellers from malaria-free areas that visit the tropics.

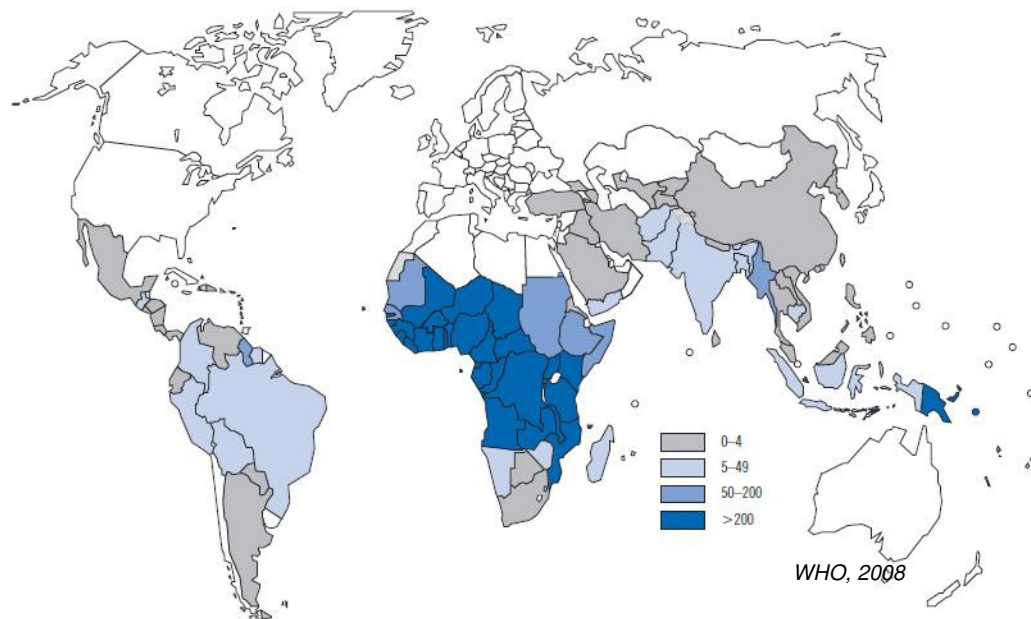


Figure 1.1 The distribution of malaria worldwide.
Estimated incidence of infection per 1000 habitants.

1.1.2. The parasite, vector and life cycle

Malaria is caused by protozoan parasites from the genus *Plasmodium* that infect and destroy the host's red blood cells (RBCs). Today four species of human parasites are known; *P.falciparum*, *P.vivax*, *P.ovale* and *P.malariae*. However, there is growing recognition that *P.knowlesi* – a simian parasite that can cause human infection – should be added to this list (Cox-Singh et al., 2008). Also, it is now proposed that *P.ovale* also represents two distinct species (Sutherland et al., 2010). Most malaria research concentrates on *P.falciparum* because it is responsible for 80 % of malaria cases and almost 90% of deaths (WHO, 2008) although there is a growing awareness of importance of *P.vivax*, whose contribution to malaria morbidity was previously somewhat neglected (Price et al., 2007; Galinski and Barnwell, 2008).

Parasites are transmitted by the female *Anopheles* mosquitoes. Various mosquito species are found to be dominant vectors in different countries, depending on the region and environment. Three African species (*Anopheles gambiae*, *Anopheles arabiensis* and *Anopheles funestus*) are responsible for the bulk of the transmission, around 30 additional ones are highly competent vectors and many others are able to transmit malaria given the opportunity (Kiszewski et al., 2004).

Plasmodium parasites are characterized by a complicated developmental cycle involving many different stages in mosquito and in human host (Fig. 1.2).

After an infectious mosquito bite, the haploid sporozoites (infectious parasite forms) migrate to the liver where they infect hepatocytes and undergo asexual reproduction (Prudencio et al., 2006). After an incubation period (which usually lasts about 10 days in *P.falciparum*) they destroy the hepatocytes and emerge from the liver as merozoites – a free stage able to infect the RBC.

Inside the red cells, a merozoite transforms into a ring stage, characterized by small size and low metabolic activity. Later a ring becomes a trophozoite - a metabolically active parasite which digests the host cell protein (consisting mainly of haemoglobin (Rosenthal and Meshnick, 1996)). The trophozoite grows and starts to undergo multiple nuclear division and segmentation to produce a schizont. Finally the schizont bursts destroying the red blood cell and giving rise to the new

merozoites, which are free to invade uninfected erythrocytes. These regular, synchronized cycles of asexual replication repeats itself every 48 or 72 hours (depending on the parasites species) until successful intervention of the immune system or the death of the host, and is responsible for most of the symptoms associated with *P. falciparum* malaria.

During this asexual erythrocytic cycle, a small number of ring parasites will transform into sexual forms – male and female gametocytes (Talman et al., 2004). When these are taken up by a mosquito in a blood meal, the gametes emerge from the red cells and fertilize to form diploid zygotes which undergo meiotic division forming new tetraploid cells. They develop into motile forms called ookinetes and penetrate the mosquito gut forming oocysts on the outer surface of the gut wall (Boete, 2005). There, in the process of multiple nuclear divisions, large numbers of haploid sporozoites are produced. After several days, the oocyst ruptures, releasing sporozoites to migrate to the salivary glands. There, the parasites become ready to infect another human host when mosquito takes next takes a blood meal, thus closing the developmental cycle.

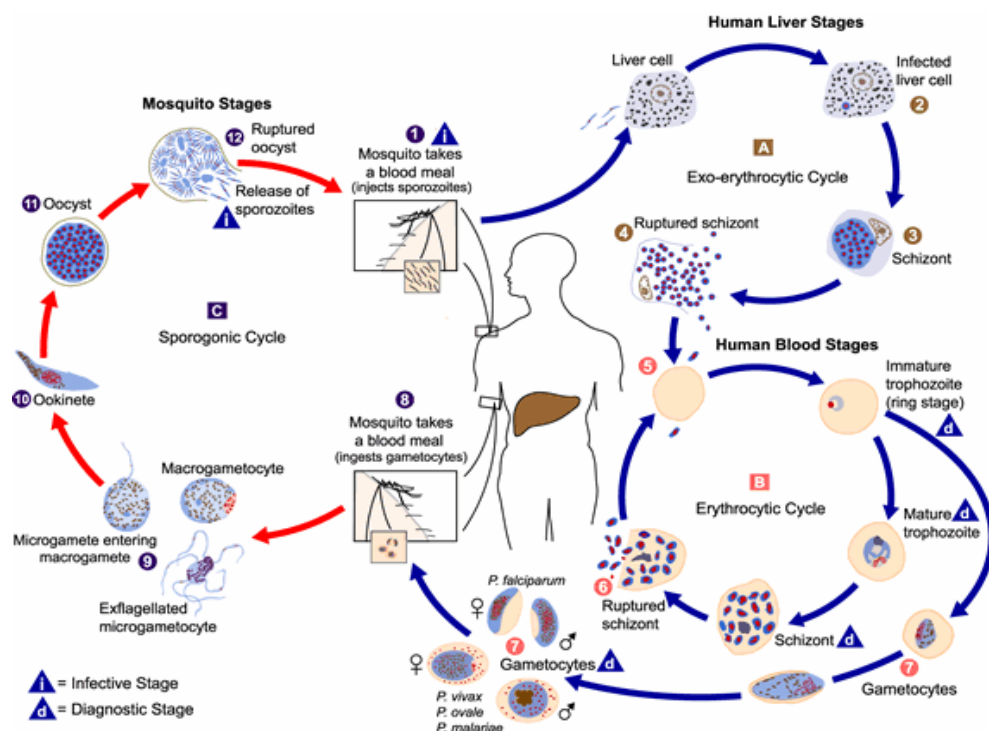


Figure 1.2 The life cycle of *P. falciparum*.

Blue arrows represent the part of the cycle in the human host, red one in the mosquito vector. The illustration from http://www.dpd.cdc.gov/dpdx/HTML/Image_Library.htm.

In *P.vivax* and *P.ovale* an additional stage is observed. After the hepatocyte infection some parasites do not develop into merozoites but into the dormant form called hypnozoites. They remain in the liver for years and are responsible for new attacks of the disease when reactivated (Cogswell, 1992).

1.1.3. Clinical picture of disease

The manifestations of infection vary from one patient to another depending on parasite strain, age of host, host immunity, general health condition etc. In some cases (for example, in older, partially immune individuals in areas of high transmission) the infection may be completely asymptomatic. Otherwise malaria may manifest itself by headache, chills, joint pains and fever. Periodic chills and fever are characteristic, recurring at intervals of 2-3 days depending on the species of parasite (they are caused by the cyclic bursting of erythrocytes freeing merozoites).

In some cases severe malaria occurs. In addition to typical malaria symptoms one or more of the following can be observed: hyperanaemia, renal failure, metabolic acidosis, pulmonary oedema or hypoglycaemia (Trampuz et al., 2003). Cerebral malaria is the most serious sub-class of severe disease and is characterized by changes in mental status followed by coma. All cases of severe malaria require prompt treatment as the untreated mortality rates are high (untreated, cerebral malaria is fatal in 90% of the cases (WHO, 2006).

Malaria can be easily misdiagnosed if the symptoms vary from the classical picture (Kallander et al., 2004). The reverse is also true – in areas of high transmission every fever tends to be interpreted as malaria, leading to possible over-diagnosis (Reyburn et al., 2004; Gwer et al., 2007). Both confer serious health consequences, as they result in a: lack of appropriate treatment for the underlying fever cause and over-use of antimalarials, encouraging the development of the resistance. (Amexo et al., 2004; Lubell et al., 2007; Skarbinski et al., 2009).

1.1.4. Disease control strategies

Today many different strategies are used in order to prevent morbidity and mortality due to malaria.

1.1.4.1.Prevention of infection

Simple means like encouraging protective clothing and closing the windows overnight can reduce the availability of humans to mosquitoes. As the mosquito's attack takes place mainly during the night, sleeping under bed nets is a particularly efficient way of protecting the local population (Lengeler and Snow, 1996).

1.1.4.2.Vector control

As the parasite is absolutely dependent on its vector, mosquito control is an essential part of malaria management programs. So far, two strategies were employed on a massive scale; indoor residual spraying (the walls of dwellings are covered with diluted insecticide that kills or repels the insects) and insecticide treated bed nets. Both achieved high success rates in reducing *Plasmodium* transmission (Beier et al., 2008; Enayati and Hemingway, 2010). However, insecticides may put evolutionary pressure on mosquitoes to develop resistance to the insecticide used (Sina and Aultman, 2001). For example the resistance to DDT – the most widely used insecticide – is currently widespread among the malaria vectors (Ranson et al., 2000).

Recently new methods of vector control have been suggested, that aim to manipulate the *Anopheles* population (so that it is less likely to transmit malaria) without eliminating mosquitoes from the environment. For instance, malaria resistant mosquitoes may be released into the wild to breed with the local populations, spreading the genes that cause them to be refractory towards malaria parasites (Marshall and Taylor, 2009). An alternative approach is to use methods that either selectively target the older mosquitoes or cause delayed death days after exposure. As mosquitoes need only 2-4 days to lay eggs after each blood feed and 10-14 days to become infectious after contracting malaria parasites, a window exists in which killing the mosquito would have only moderate effect on its reproductive

success but would stop malaria transmission (Read et al., 2009). The use of fungal biopesticides (Blanford et al., 2005), genetically engineered *Wolbachia* bacteria (Read and Thomas, 2009) and densoviruses were proposed.

1.1.4.3.Vaccines

Unfortunately, despite huge investment, no fully effective vaccine for malaria has been developed so far, although there is a pipeline of candidates currently entering or undergoing the various stages required for clinical testing (Girard et al., 2007). Previous attempts have resulted in limited success, possibly because of incomplete understanding of critical targets, variation in parasite antigens, poor development of natural lifelong immunity even after repeated exposure (suggesting that immune memory is not stimulated well) or parasite evasion of the immune system. Therefore new and innovative methods may be needed to elicit the appropriate immune response (Pierce and Miller, 2009; Vanderberg, 2009).

1.1.4.4.Chemotherapy

In practice, much of the control of disease involves the use of chemotherapy – either treatment of infected patients, or prophylaxis. It is described in detail in the next section.

1.2. Malaria chemotherapy

Chemotherapy remains the key tool for malaria control reducing mortality and morbidity as well as decreasing the parasite transmission worldwide. A variety of antimalarials exists. Some are derived from natural herbal remedies (like the bark of the cinchona tree in South America and sweet wormwood plant in China); others are the results of *de novo* synthesis or modification of existing compounds.

Unfortunately, use of these compounds has often resulted in the evolution of drug resistance, eventually leading to reduced clinical effectiveness.

This section reviews some of the most important classes of antimalarials, with particular focus on chloroquine and artemisinins that are the subject of this thesis. Section 3 focuses on different aspects of resistance to antimalarials.

1.2.1. Antifolates

Antifolates are drugs that affect the synthesis and utilization of folate. A supply of tetrahydrofolate is required for the synthesis of DNA and some amino acids. It is particularly important in metabolically active, rapidly dividing cells such as malaria parasites during their asexual cycle. In *Plasmodium* folate may come principally from *de novo* synthesis because parasites may not be very efficient in salvaging the host folates (Hyde, 2005). The folate pathway appears to be a good target for antimalarial drugs: its successful inhibition leads to the cell cycle arrest and death.

The antifolates used in malaria treatments are divided into two groups, often called class I and class II antifolates, depending on their target (Nzila, 2006). The class II antifolates inhibit dihydrofolate reductase (*dhfr*) - an enzyme catalysing one of the final steps of folate pathway. The most popular drugs in this group are pyrimethamine (currently discarded in many places because of resistance), proguanil (often used for prophylaxis), and chlorproguanil. The drugs from the second group (class I antifolates) are inhibiting dihydropteroate synthase (*dhps*) an enzyme further upstream in the pathway. Sulfadoxine and dapson are the most common examples. The class II antifolates are rarely used alone. In malaria treatment the class I and II

antifolates are often combined for better effect. The most common combinations include pyrimethamine/sulfadoxine (Fansidar), pyrimethamine/sulfalene (Metakelfin), proguanil/dapsone and chlorproguanil/dapsone (Lapdap).

1.2.2. Aminoquinolines and chloroquine (CQ)

This wide group of antimalarials includes a number of very effective commonly used drugs as well as some historically important ones. They are characterized by an aromatic quinoline skeleton with a modified amino group attached to it. They are divided into two main subgroups based on the chemical structure of the compounds. The 4-aminoquinolines (*e.g.* chloroquine, amodiaquine), are powerful blood schizontocides, killing mainly the parasite stages carrying out haemoglobin digestion. 8-aminoquinolines (primaquine, pamaquine) are also effective against the pre-erythrocytic liver stages which makes them a useful tool for prophylactic treatment (as they target the early infection stages) and against non-*falciparum* malaria (*P.vivax* and *P.ovale*) in which the dormant liver stages may be the source of reinfection (relapses) if not eradicated properly.

Chloroquine is the most important and well-studied drug from this group

1.2.2.1. Chloroquine

Chloroquine (Fig. 1.3a) was first synthesised in 1934 in response to the general demand for a cheaper replacement for cinchona bark. Interestingly, initial animal tests and a small-scale clinical trial performed in 1935 suggested that it was “too toxic to be used for humans” (Coatney, 1963). However, when during the 2nd World War the leading antimalarial treatment (cinchona bark) became unavailable in Europe due to socioeconomic reasons; chloroquine (this time in different formulation) was revisited and, as the results of the repeated tests were very encouraging, officially introduced on the market in 1944. It turned out to be a perfect drug for use in tropics: very effective, cheap, easy to administer orally, stable at room temperature and safe (even during pregnancy). Unsurprisingly it quickly became the first line of antimalarial treatment worldwide for over 20 years. However, resistance began to develop in Southeast Asia and South America, spreading to Africa by 1978. By 2002 the WHO officially recommended the withdrawal of chloroquine due to widespread resistance and in most countries it is currently replaced with alternative

therapies. However, it is still used (mostly in Africa), due to its low price and lack of access to the other therapies.

1.2.2.2.The mechanism of action

Similarly to other 4- aminoquinolines chloroquine is a schizontoid targeting the asexual blood stage of malaria and it does not affect liver stages (Belnoue et al., 2004) or gametocytes (Chevalley et al., 2010). Chloroquine acts in the parasite digestive vacuole – the lysosomal-like compartment characterised by a very low pH in which host protein is digested by proteolytic enzymes (Saliba et al., 1998). Here, host haemoglobin is digested by proteases releasing amino-acids and the porphyrin ring with co-ordinated iron in the centre – hem (Goldberg and Slater, 1992). In its free form hem is a highly active compound, capable of generating a cascade of free radicals and thus is toxic to the parasite. Therefore in the digestive vacuole hem is polymerised into the crystals of inactive hemozoin (Slater et al., 1991). The interference with this mechanism is universally considered to be a main pathway of CQ action (Slater and Cerami, 1992).

Chloroquine is a weak base that diffuses freely into the digestive vacuole (Fig. 1.3a) where it gets diprotonated and, as the vacuolar membrane is impermeable to charged compounds, trapped within the compartment. There it binds to ferriprotoporphyrin IX (FPIX, oxidised hem derivate) (Jearnpipatkul et al., 1980) disturbing the formation of hemozoin and keeping FPIX in its active form. This results in the accumulation of free FPIX and/or CQ-FPIX complexes, which damage the parasites, possibly by oxidative stress (radical oxygen species), ultimately leading to its death (Fig. 1.3b).

The mechanism described above is considered to be the main pathway of chloroquine action. However some other reports suggest that the drug is effective also against the parasite stages (like young rings and schizonts) that are not involved in hem metabolism and do not possess (in the case of ring stages) a fully formed digestive vacuole (Gligorijevic et al., 2008; Sharrock et al., 2008). Moreover CQ may also decrease multiplicity in nuclear divisions observed in schizonts (suggesting interference with DNA replication) and cause a delayed death syndrome (characteristic for some antibiotics interfering with apicoplast function (Dahl and

Rosenthal, 2007)) (Gligorijevic et al., 2008). These data suggest the possible existence of alternative pathway by which chloroquine can kill the parasite.

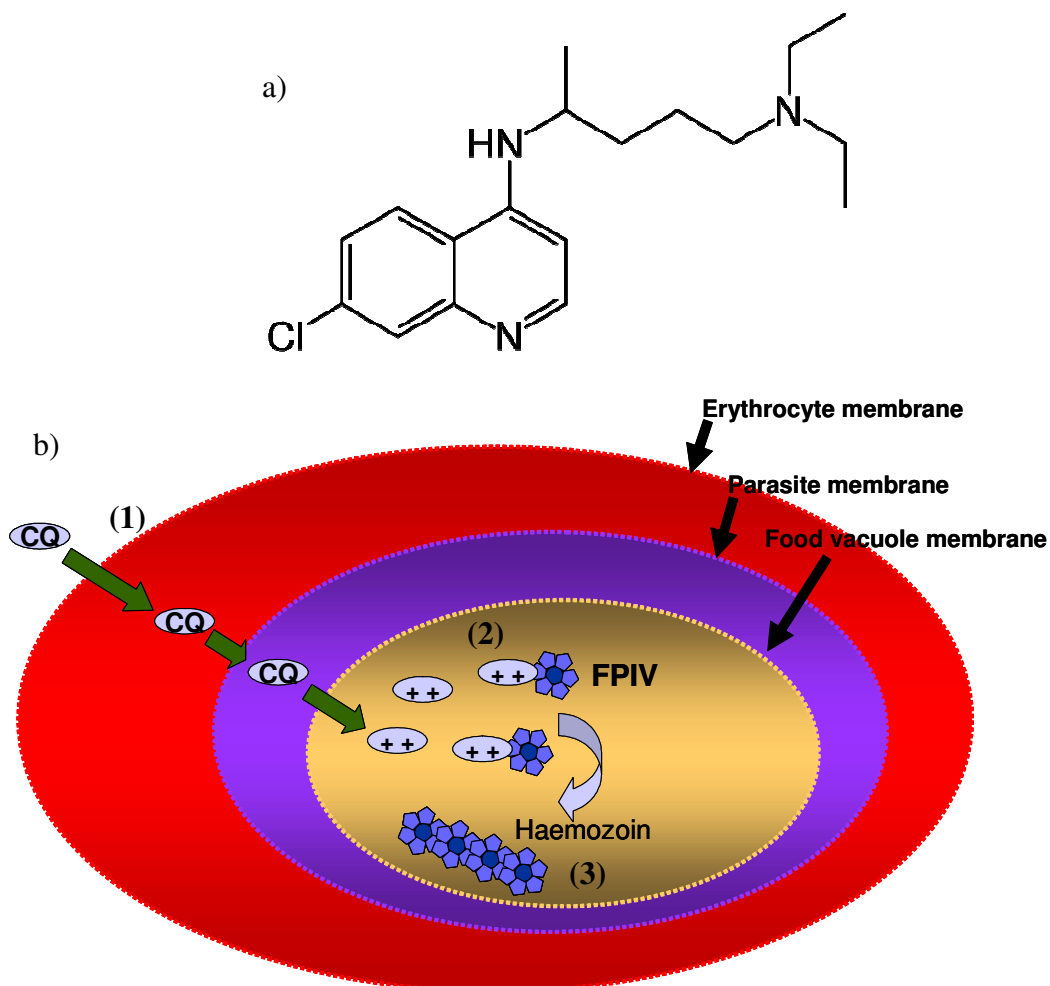


Figure 1.3 The chemical structure of CQ and the putative mechanism of its action.

a) the chemical structure of CQ, b) the putative mechanism of CQ action. The drug is diffusing into the digestive vacuole (1) when it gets protonated and binds to hem (2), inhibiting the formation of hemozoin (3)

1.2.3. 4-methanolquinolines

Very similar to the previous group, the 4-methanolquinolines are characterised by the same aromatic skeleton, this time however the amino group is replaced with the methanol one. The two important drugs from this group are quinine – a naturally occurring alkaloid isolated from the cinchona tree bark, one of the earliest malaria remedies - and mefloquine, a long acting synthetic antimalarial, now more widely used as part of combination therapies. The mechanisms of action of 4-methanolquinolines is not known although some physiological effects of these drugs suggests that they also intervene with some stage of haemoglobin catabolism (Famin and Ginsburg, 2002).

1.2.4. Artemisinin

The antimalarial properties of *Artemisia annua* or Sweet Wormwood, a common plant across central Asia, have been known since ancient times in Chinese medicine. In 1972 an active antimalarial compound - artemisinin - was isolated and derivatives with better pharmaceutical proprieties were soon synthesised. All artemisinins are sesquiterpene lactones containing an endoperoxide bridge in their structure. They are all very short lived (no greater than about 2 hours half-life) and very effective not only against blood stages (Adjuik et al., 2004) but also against the young gametocytes (thus interrupting the parasite transmission) (Pukrittayakamee et al., 2004).

1.2.4.1. Artemisinin combination therapy (ACT)

Currently artemisinins (mainly dihydroartemisinin, artesunate and artemether) are recommended as the first line of treatment worldwide. As WHO strongly discourages the use of artemisinins as monotherapy (in order to delay the appearance of resistance (White and Olliaro, 1996)), they are usually recommended to be combined with other drugs, usually long-lasting ones, and with a different mechanism of action. The most common partner drugs are mefloquine (Lariam[®]), lumefantrine, amodiaquine, piperaquine and antifolates. ACTs are the only form of therapy with universal efficiency towards drug resistant strains.

1.2.4.2. The mechanism of action

The mechanisms of action of artemisinins are less understood than those of chloroquine. The general mechanism of molecule activation is known - all artemisinins contain a peroxide bridge which, when cleaved (probably in the presence of iron Fe^{2+}), generates a cascade of free radicals within the cell (Meshnick et al., 1996a). However, the object of this attack remains to be identified. The proposed targets included hem (Hong et al., 1994; Meshnick, 1994), PfATP6 (sarco/endoplasmic reticulum membrane calcium ATP-ase) (Eckstein-Ludwig et al., 2003; Jung et al., 2005), alkylation of specific proteins (Yang et al., 1993; Meshnick, 1994) and, recently, the mitochondrial oxidative chain (Li et al., 2005; Wang et al., 2010). Most likely, as in the case of chloroquine, it is a complex interplay between multiple pathways that ultimately leads to parasite death.

1.2.5. Atovaquone

Atovaquone is a synthetic drug that is effective against a range of different protozoan parasites. It is found to target the cytochrome *bc1* complex, disrupting mitochondrial electron transport and thus interfering with ATP synthesis (Mather et al., 2005). In malaria treatment it is used in combination (Malarone[®]) with proguanil with which it seems to be sharing some synergistic effects (Srivastava and Vaidya, 1999).

1.2.6. Antibiotics

Tetracycline, doxycycline, clindamycin and azithromycin can all be used (usually as part of ACT) for malaria treatment (WHO, 2006; Noedl, 2009). The broad spectrum of action of antibiotics can be useful as they can also treat the additional infections often accompanying malaria, or (in cases of misdiagnosis) the true cause of the symptoms interpreted as malaria. However for the same reason they are not very suitable for prophylaxis and self-treatment and are more often used as the second-line treatment for infections that don't respond well to traditional antimalarials (WHO, 2006).

1.3. Drug resistance in malaria

In the second half of 20th century the growing efficiency of insecticides and antimalarials and improved standards of medical care prompted the WHO to predict that malaria would be eradicated by the end of the century (Jeffery, 1976). Today malaria seems far from eradication and its mortality is still high. The increasing frequency of drug resistant parasites is considered one of the major factors contributing to this situation. Currently, parasite strains resistant to all major classes of drugs are present in the field and resistance to some of them so widespread that they have been withdrawn from use.

1.3.1. Assessing resistance

How do we know that resistance to a given drug has appeared in a population? Today local monitoring networks in many countries concentrate on monitoring the resistance status of local parasite populations. Three different approaches can be used: field studies, *in vitro* testing and verification of molecular markers of resistance.

1.3.1.1. *In vivo* – field studies

During field studies the reaction of patients to a treatment is recorded and resistance is suspected if increased rates of treatment failure are observed. To this end, treatment failure is defined as incapability to clear parasitaemia (parasitological failure) and/or resolve clinical symptoms (clinical failure) despite the administration of the drug (WHO, 2009). A treatment failure, however, does not always imply resistance - it can be caused by inadequate drug absorption, poor adherence to the treatment etc. The reverse is also true - the adequate clinical response can be observed with the resistant strain if the individual is semi-immune. Therefore field data may need to be confirmed by laboratory testing of culture adapted parasites.

1.3.1.2. *In vitro* - laboratory testing

Compared to field studies, laboratory tests are more accurate in defining and quantifying parasite-determined resistance. In laboratory conditions cultured parasites can be exposed to precisely known concentrations of drug and the degree of

growth inhibition can be accurately recorded. The resistance is usually defined as a shift in dose response curve resulting in change of IC_{50} (a concentration of the drug required to kill 50% of parasites) or more rarely IC_{90} (see Fig. 1.4). The *in vitro* test are not used for routine surveillance, however they can be useful to resolve many aspects of drug resistance. It includes the validation of molecular markers, detection of cross resistance patterns and monitoring the resistance to drugs used in combination therapy (when the treatment failure would require the simultaneous resistance to both of them).

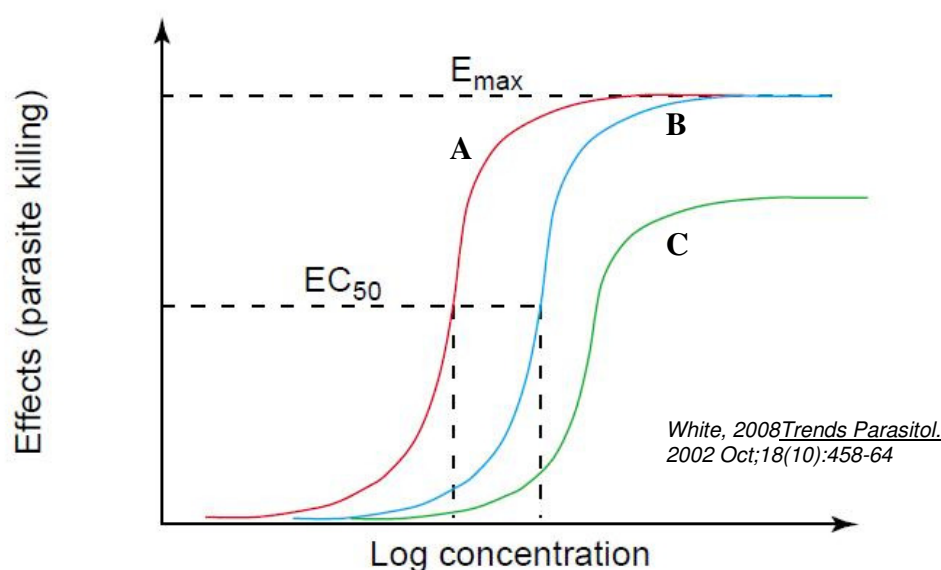


Figure 1.4 Resistance observed *in vitro*

The difference in dose response curve between the sensitive (a) and resistant (b,c) parasites. The resistance is achieved either by the shift of the response curve (b) or change of its shape (c). Figure according to (White, 2002).

1.3.1.3. Use of molecular markers

Once the molecular markers of resistance are identified, they can be used to monitor the resistance in field. They can detect even the small prevalence of a resistant strain present in a population, which is useful, especially when mixed infections are common and a resistant strain can be easily masked by the sensitive population. Moreover they tend to be quicker and more reliable than other tests. Due to their price they are rarely used for routine surveillance. However they are

replacing *in vitro* tests in many specific uses including monitoring for the first signs of emerging resistance in a given population, following-up of the fate of a resistant allele once the drug has been withdrawn, monitoring resistance to one of the compounds in a combination therapy etc.

1.3.2. Evolution of resistance

1.3.2.1. Appearance and spread and resistance - how easy it is for parasite?

From an evolutionary point of view, the rise and spread of resistant parasites in a population will depend on two factors – the probability of an initial mutational event and the reproductive advantage it confers to the parasite. In theory both of those should be very high. On the one hand the sheer numbers of parasites ensure the constant generation of large number of mutations. An infected adult can easily carry 10^{12} parasites (which corresponds to ~3% of parasitaemia) and the new generation appears with each round of asexual replication. Combined with the mutation rate ($\sim 10^{-9}$ /base/generation (Conway, 2007)) it means that, if the resistance is conferred by single SNP mutation, $\sim 10^3$ resistant parasites will arise within each patient. On the other hand a drug intervention constitutes a very strong selection pressure and current massive drug use ensures that thousands of infections are exposed to the treatment every day. Therefore the resistance would appear fairly often and we should observe multiple origins of resistance mutations.

The field data however suggests that resistance arises from few mutational events that spread throughout populations rather than from many multiple, independent origins. This is true for three of the most widely spread resistances: chloroquine (Wootton et al., 2002), sulfadoxine-pyrimethamine (Nair et al., 2003) and (to a lesser extent) mefloquine (Nair et al., 2007).

This may be explained by an initial bottleneck in the infection. Once a resistant parasite appears, it has to successfully compete with other parasites present in the same host, achieve a significant proportion of infection and produce gametocytes that are going to be ingested by mosquitoes and transmitted to another patient. Then the whole cycle is repeated. The probability of all these events is very low and is fairly independent of the selective advantage conferred by resistance - even very fit genotypes can appear late in the infection and fail to produce

gametocytes before a mosquito bite or end up in a partially immune host that is going to clear the infection regardless of its resistant status. Only after passing this crucial stage does resistance spread according to the reproductive advantage that it confers to the resistant parasite. If we compare it with the other potential obstacles (necessity of competing with other genetically different strains in mixed infections, potential fitness cost of resistance (see section 3.3), *var* gene switching etc.) we obtain actual mutation rates that may be many orders of magnitudes lower (Hastings, 2004). It helps explain why we observe so few origins of mutations.

The initial obstacles may also explain why the resistance to many antimalarials seems to arise in the same geographical regions – the resistance “hot-spots” characterized by low transmission and low level of immunity in the population (like South-east Asia (most notably Thai-Cambodian border) and equatorial South America). In that setting the majority of infections are treated (asymptomatic malaria is rare in a non-immune host) and the mixed infections in which parasites will have to compete against a different sensitive strain (possibly without drug treatment) are less frequent. Therefore a resistant allele is less likely to be eliminated before reaching the critical frequency allowing the subsequent spread.

1.3.2.2. Use of combination therapy to slow the development of resistance

The use of combination therapy is an example of how the evolutionary forces responsible for the development and spread of the resistance can be used in order to extend the useful lifespan of a drug. It was successfully used in tuberculosis and HIV treatment and now it is applied to malaria. The concept is simple - in order to slow the evolution of resistance and extend the useful lifetime of drugs, two or more antimalarials are paired together. If resistance to one drug is hard to develop the probability of acquiring resistance to two drugs simultaneously is very low (possibly next to 0) (White and Olliaro, 1996).

Currently used ACT are usually composed from a short acting artemisinin (ART) derivative paired with a long lasting compound of different mechanism of action (e.g. mefloquine, amadioquine, antifolates). This combination has some advantages. ART compound is a powerful schizonticide, quickly reducing most of the parasite burden. The second drug instead, remains in the patient blood long after,

killing the remaining parasites and protecting from re-infection, thus reducing the frequency of survival of parasites exposed to subcurative concentration of ART, (which encourages resistance). Some argue however that this mismatch in the half-lives of the drugs is not ideal as it encourages the spread of the resistance to the long-acting drug as, in the absence of ART, it functions as it would in monotherapy.

Moreover, some of the partner drugs were previously used as monotherapy and resistance to them is already at least partly established. In that case ACT would lead to further spread of resistance to partner drugs (although it would still delay the emergence of resistance to artemisinin compared to monotherapy). Therefore the half-lives of two ACT components should be similar, to slow down development of the resistance towards both compounds.

1.3.3. Fitness cost of drug resistance

Already Charles Darwin in “The origin of species” suggested that the development of one trait in the population may come at the expense of another (Darwin, 1859). The trait that is an advantage in one environment can be a burden to another. The same applies to drug resistance - the mutations that allow the organism to tolerate the higher level of curative agent are likely to reduce its performance in the drug free environment (Fig. 1.5). This phenomenon called “fitness cost” has been observed in such different organisms as *Mycobacterium tuberculosis* (Burgos et al., 2003), Human immunodeficiency virus (HIV) (Borman et al., 1996) and *Staphylococcus aureus* (Besier et al., 2008). Now an accumulating body of evidence suggests that it may also apply to the resistance to antimalarials.

1.3.3.1. Evidence from field

If a fitness cost of resistance exists, after the withdrawal of the drug pressure, the decline in frequency of resistant alleles in the population should be observed because of their gradual replacement with fitter, more sensitive alleles. Several field studies reported such decreases following the withdrawal of pyrimethamine (Clyde and Shute, 1959; Zhou et al., 2008) or chloroquine (Thaithong et al., 1988; Mita et al., 2003; Temu et al., 2006). The most spectacular example comes from Malawi. After replacing chloroquine with alternative therapy, the frequency of resistant allele decreased from 85% in 1992 to 13% in 2000 (Kublin et al., 2003; Mita et al., 2003;

Laufer et al., 2006) and a recently performed clinical trial shown that CQ is again fully effective in this region (Laufer et al., 2006).

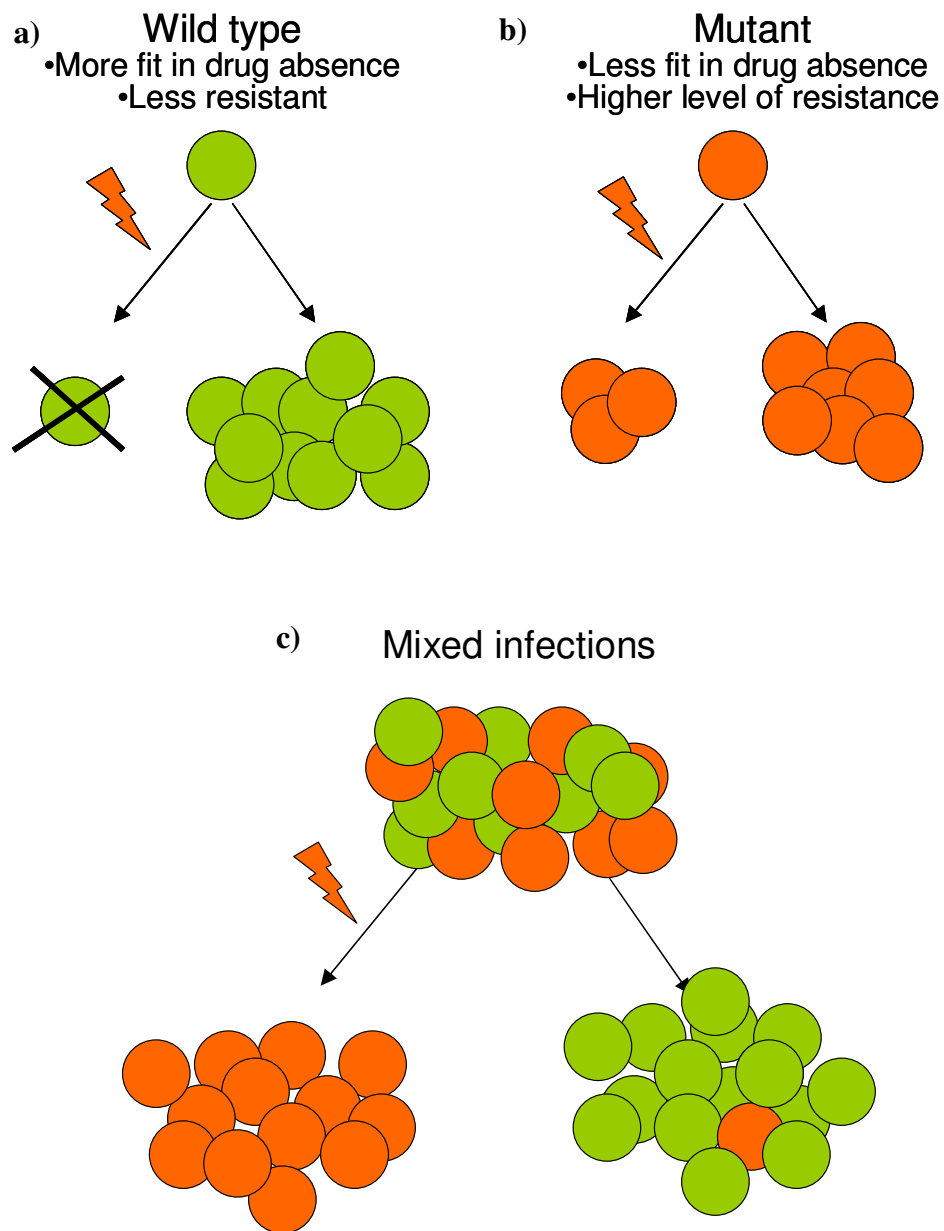


Figure 1.5 Fitness cost of drug resistance.

The green and red circles represent sensitive and resistant parasites respectively and ⚡ sign represents the drug action. The sensitive mutant (a) is inhibited by the drug but grows poorly in drug free conditions. The resistant mutant (b) can survive the drug treatment but its growth is handicapped. Therefore in mixed infection the resistant parasite will dominate but under lack of treatment the sensitive parasite is going to outgrow the mutant

Additionally, in some countries, characterised by variable transmission rates between the wet and dry season, seasonal variation in the frequency of resistant alleles was observed. During the wet season (when transmission is high and more infections are symptomatic and treated) the frequency of resistant allele was increasing but during the dry season (when the parasite persists mainly in asymptomatic untreated infections) this frequency decreased (Abdel-Muhsin et al., 2004; Babiker et al., 2005; Ord et al., 2007). These studies also suggest the existence of fitness costs associated with resistance (Babiker et al., 2009)

1.3.3.2. Evidence from laboratory studies

In the laboratory two strains of different levels of resistance can be competed in controlled conditions. Ideally the two isogenic strains: resistant and sensitive one are used. This can be achieved by genetic modification that replaces a resistant allele with a sensitive one or *vice versa*, or by laboratory selection of a resistant strain and its comparison with its sensitive progenitor. So far, two such experiments were performed in *P.falciparum*. They were studying the impact of *pfmdr1* alleles associated with resistance to aminoquinolines (Hayward et al., 2005) and the mutations in cytochrome *c* conferring atovaquone resistance (Peters et al., 2002). Both experiments suggested the existence of fitness cost.

Alternatively the rodent model was used when a laboratory selected resistant strain can be compared with its sensitive progenitor. For example the analysis of parasites cloned out of the mixture of resistant and sensitive strains at various points of the infection revealed that pyrimethamine resistant *P.chabaudi* (Rosario et al., 1978; Peters et al., 2002; Walliker et al., 2005) has reduced growth when competed with its sensitive progenitors. Moreover the analysis of the cross between the Pyr-resistant and sensitive parasites revealed that the causative mutation in *dhfr* gene conferring the resistance is gradually eliminated from the population, providing the conclusive evidence that the fitness cost is due to the locus of pyrimethamine resistance (Rosario et al., 1978). However, it has to be added that in *P.chabaudi* pyrimethamine resistant parasites, a fitness cost was observed only initially: after further passage a clone had recovered its original fitness, possibly because of the appearance of supplementary, compensatory mutations (Walliker et al., 2005).

P. chabaudi was also used in order to investigate chloroquine resistance. Again the parasites cloned at various time point from the mixed infection were analysed, however in this case no evidence of fitness cost was discovered (Rosario et al., 1978).

1.3.3.3. Consequences

What impact can the fitness cost have on the spread of resistant alleles in the population?

First of all it can slow down the spread of the resistance. If the competition of resistant and sensitive strain in untreated hosts results in the domination of sensitive alleles, the likelihood of an accidental elimination of resistance mutations during the initial stages of resistance development, described in section 3.2 is greater.

Secondly, the fitness cost can potentially increase the success of combination therapy – if the presence of second drug decreases the evolutionary advantages of resistance in a treated infection and the fitness cost decreases its survival of resistant parasites in an untreated one, the net result can lead to elimination of the resistance allele (Hastings and Donnelly, 2005).

Finally, once the drug is withdrawn, the fitness cost may cause the elimination of resistant alleles from population making possible at least temporary re-introduction of the drug. However this process would most likely take a long time and, as resistant alleles would probably still be present in very low frequency, the resistance would probably re-emerge rapidly after the drug reintroduction. It remains to be investigated whether some strategies (like using multiple treatments regimes or temporary drug switching) could be potentially used to keep the resistance at the low level.

1.4. Chloroquine (CQ) resistance

1.4.1. Appearance and spread of CQ resistance

Chloroquine resistance appeared after almost 20 years of heavy use of the drug. The first reports of treatment failures arrived simultaneously from South America (Moore and Lanier, 1961) and Southeast Asia (Harinasuta et al., 1965). From there the resistance spread across those continents before reaching Africa in the 1970's - the first cases of chloroquine resistant *P.falciparum* were reported in Tanzania in 1979 (Campbell et al., 1979). After that the resistant parasite spread across Africa. Soon it was estimated that 80% of world parasite population exhibited some level of CQ resistance (Ginsburg, 2005) and in many countries it reached fixation.

The rise and spread of CQ resistance is considered to be one of the important factors contributing to increase of malaria deaths in Africa in the last 30 years of the 20th century (Grech et al., 2002).

1.4.2. Genetics of CQ resistance

The phenotyping of progeny of the cross between CQ-resistant and CQ-sensitive strains of rodent malaria parasite *P.chabaudi* revealed that CQ resistance was inherited as a simple Mendelian trait and therefore was likely to have a genetic basis (Rosario, 1976). However, the main determinants of the resistance remained elusive and the lack of necessary tools (like the *in vitro* culture system or polymerase chain reaction) made the investigation very difficult. Only in the 1990's the isolation of a CQ-resistant clone (~Dd2) and advances in molecular biology techniques made possible the genetic analysis that revealed the main resistance determinants in the major human parasite *P.falciparum*.

1.4.2.1. *Pfcr*

In 1990 the cross between two *P.falciparum* clones: chloroquine-sensitive (HB3) and resistant (Dd2) was performed (Wellems et al., 1990). The genetic analysis of progeny from this cross indicated a locus on chromosome 7 as the main determinant of the resistance (Wellems et al., 1991). Later the analysis of the additional progeny has narrowed this locus to 38 kb fragment containing only 10 genes (Su et al., 1997). After ruling out *cg2* gene (initially the prime candidate), resistance was mapped to the highly polymorphic ~3.1 kb gene composed from 13 exons (Fidock et al., 2000). It encoded a 424 amino acid long 48.6 kDa protein, forming 10 transmembrane domains situated in the parasite digestive vacuole (Fig. 1.6). Due to the transporter-like features of its secondary structure the new protein was called *P.falciparum chloroquine resistance transporter - PFCRT*. Its original function is unknown, however bioinformatic analysis (Martin and Kirk, 2004; Tran and Saier, 2004) revealed that it is a part of drug-metabolite transporter superfamily and Martin and Kirk suggest that in its native state *pfcr* is a transporter of small cellular metabolites from the parasite vacuole, possibly peptides or amino acids resulting from haemoglobin digestion.

The sequencing of *pfcr* in isolates from different parts of the world indicated the strong correlation between the mutations within this gene and CQ resistance (Fidock et al., 2000; Chen et al., 2003; Nagesha et al., 2003; Durrand et al., 2004). Moreover microsatellite analysis across the *P.falciparum* genome revealed a selective sweep around *pfcr* (Wootton et al., 2002) (A selective sweep is a decrease in genetic variation in the particular region, usually caused by recent intensive evolutionary pressure (in this particular case - the massive CQ use)). The final proof for the connection between *pfcr* mutations and CQ resistance came from a series of genetic modification (transfection) experiments. The introduction of a resistant *pfcr* allele into the sensitive parasite resulted in the significant increase of IC₅₀ for chloroquine (Fidock et al., 2000; Sidhu et al., 2002) thus establishing *pfcr* as a genetic marker of resistance.

So far, 15 different mutations have been discovered in resistant alleles of *pfcr*. (see Fig.1.6 for the exact locations). The most important of those - the K76T substitution appears to be crucial and present in all resistant strains. The allelic exchange

experiments proved that the removal of this mutation is sufficient to convert the CQ resistant allele into sensitive one (Lakshmanan et al., 2005). Interestingly, the K76T mutation was never observed alone; all identified resistant *pfcr*t alleles contain 3-9 additional substitutions. The pattern of these additional mutations is characteristic for geographic region (Fidock et al., 2000). For example the sequence of residues 72-76 (CVMNK in sensitive parasite) is mutated into CVIET in most of African and south-east Asian resistant strains. In South American ones, instead, the SVMNT haplotype is prevalent.

The further analysis revealed that the K76T mutation appeared independently only around six times worldwide and each time was connected with slightly different set of secondary/pre-existing mutations (Mita et al., 2009). These additional mutations do not seem to confer in themselves CQ resistance – parasites containing the “resistant” *pfcr*t allele without K76T mutation were found to be fully sensitive. However their presence seem to be necessary to obtain a fully functional *pfcr*t allele: no parasite carrying K76T mutation alone was identified in field and the attempts to generate such parasite in laboratory failed because of extremely low replication rate of the obtained clone (Lakshmanan et al., 2005). Therefore it is possible that main role of the additional mutations is to compensate for deleterious effects of K76T change (Valderramos and Fidock, 2006).

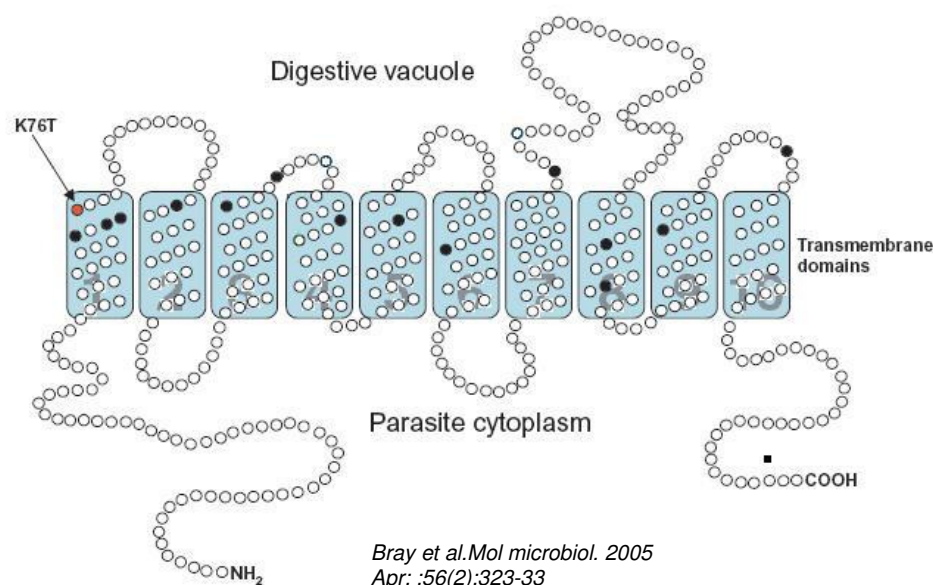


Figure 1.6 *Pfcr*t – the main determinant of CQ resistance in *P.falciparum*

The structure of *PFCRT* protein. The crucial K76T mutation is marked in red; other residues mutated in CQ-resistant alleles are marked in black. The illustration is taken from (Bray et al., 2005).

1.4.2.2. *Pfmdr1*

Before the discovery of *pfcrt* as the principle determinant of CQ resistance, it was shown that CQ resistant *P. falciparum* shares some similarities with drug resistant mammalian tumor cells; including the reduced accumulation of drug in intracellular compartment and reversal of resistance by the calcium channel blocker verapamil (Krogstad et al., 1987). In tumor cells these features are mediated by a membrane transporter called P-glycoprotein encoded by *mdr1* gene (Juliano and Ling, 1976). The search for its homologue in malaria parasite revealed *pfmdr1* – *Plasmodium falciparum* multidrug resistance gene (Foote et al., 1990). *Pfmdr1* is situated on chromosome 5 of *Plasmodium falciparum* and encodes a 162 kDa trans-membrane protein with 12 transmembrane domains organised in two symmetrical units (Fig. 1.7). The *PFMDR1* protein is a member of the ABC (ATP binding cassette) transporter superfamily and (similarly to *PFCRT*) it is localised within the digestive vacuole membrane (Cowman et al., 1991). It is thought to transport solutes (including some antimalarial drugs) into the digestive vacuole (Rohrbach et al., 2006).

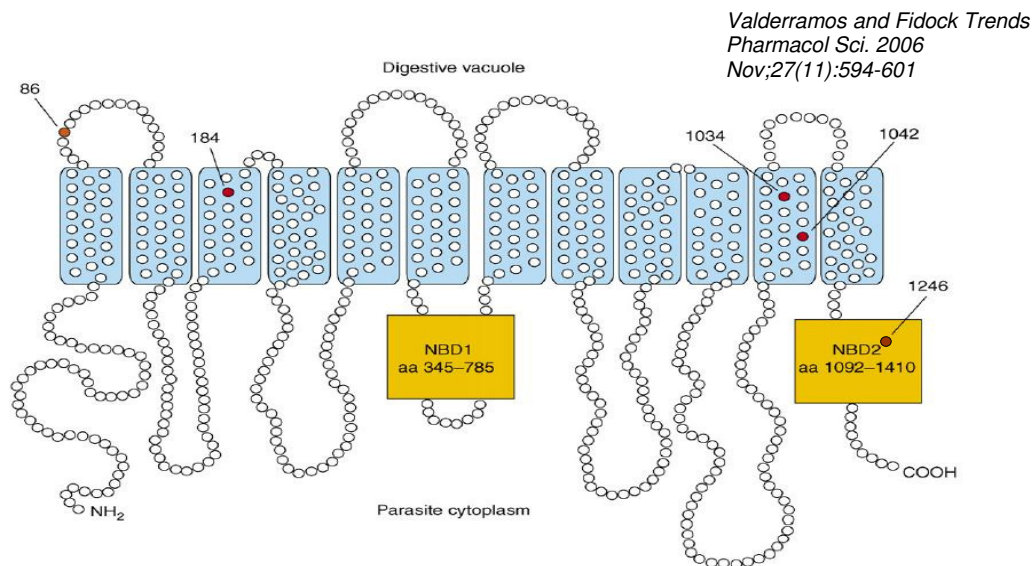


Figure 1.7 *PFMDR1* – a multidrug resistance protein

The structure of *PFMDR1* protein. Mutations associated with high CQ resistance marked in red. NBD- nucleotide binding domain. Illustration taken from (Valderramos and Fidock, 2006).

The sequencing of the *pfmdr1* gene in CQ-resistant and CQ-sensitive isolates revealed five polymorphisms: N86Y Y184F, S1034C, N1042D and D1246Y (Foote et al., 1990) potentially involved in CQ responses. The N86Y mutation showed statistically significant correlation with the high CQ resistance in many field studies (von Seidlein et al., 1997; Babiker et al., 2001; Nagesha et al., 2001). However it could not be classified as its main determinant. There were no associations with the other polymorphisms.

The role of *pfmdr1* was further elucidated in laboratory experiments. The genetic cross between the resistant and sensitive strain showed a lack of association between this gene and CQ responses of the progeny, which disqualified *pfmdr1* as the universal resistant determinant (Wellems et al., 1990). Later a series of allelic exchange experiments (Reed et al., 2000; Sidhu et al., 2005) revealed that the introduction of triple mutation (S1034C, N1042D and D1246Y) into *pfmdr1* allele of the sensitive strain does not confer in itself CQ resistance. Instead it increases the level of resistance in some parasites strains already carrying the *pfcr1* K76T mutation. This effect was however highly strain-dependent.

In summary, evidence suggests that *pfmdr1* has an epistatic role (*i.e.* its influence on the phenotype is dependent on the genetic background it is found on) and can increase the level of resistance in resistant 76T *pfcr1* strains. It also explains a non-random association between the *pfcr1* K76T and *pfmdr1* N86Y alleles observed in many endemic regions (Duraisingh et al., 2000; Adagu and Warhurst, 2001; Babiker et al., 2001; Mita et al., 2006).

Initially it has been also considered that the amplification of *pfmdr1* gene and the resulting overexpression of the protein (encountered in some of the resistant isolates) can increase the level of resistance to CQ (Foote et al., 1989). However, further investigations did not support this conclusion. In fact the prolonged exposure to high concentration of chloroquine resulted in de-amplification of this gene in a mefloquine-selected laboratory strain W2-mef, suggesting that *mdr1* amplification leads to reduced CQ resistance (Barnes et al., 1992).

Associations between *pfmdr1* mutations and duplication and the responses to other antimalarial drugs, including mefloquine, halofantrine, quinine and artemisinin

has also been found (Duraisingh and Cowman, 2005; Sidhu et al., 2005), suggesting that its action is general rather than CQ-specific.

1.4.2.3. Unresolved questions concerning genetics of chloroquine resistance

Although a strong correlation was found between CQ resistance and the *pfcr*t and *pfmdr*1 mutations, other proteins are likely to be involved. The phenotyping of field isolates (Djimde et al., 2001; Durand et al., 2001; Chen et al., 2002) as well as allelic exchange experiments (Valderramos et al., 2010) has shown that even parasites with the same allelic variants of *pfcr*t and *pfmdr*1 exhibit a wide range of possible chloroquine responses, suggesting that other genes may modulate the response of parasites to CQ. Additionally, the crucial K76T substitution is found in all of the resistant strains; however it has been also identified in some isolates with an apparently effective chloroquine response. Therefore the mutated *pfcr*t may not always lead to CQ resistance.

Most importantly however, the other major human parasite -*P. vivax*, whose contribution to malaria mortality and mortality is often underestimated (Mendis et al., 2001; Price et al., 2007; Galinski and Barnwell, 2008), can also develop CQ resistance. The reports of chloroquine-resistant *P. vivax* come mainly from South – East Asia (Phan et al., 2002; Srivastava et al., 2008) and Indonesia (Sumawinata et al., 2003; Ratcliff et al., 2007) but it can be also observed in South America (Soto et al., 2001; Ruebush et al., 2003). Both *pfcr*t and *pfmdr*1 homologues in *P. vivax* were investigated for their connection with CQ resistance. However, although a number of polymorphisms were discovered, there was no strong connection between any of them and the treatment outcome in field level of resistance (Nomura et al., 2001; Suwanarusk et al., 2007; Barnadas et al., 2008; Orjuela-Sanchez et al., 2009). Therefore the genetic basis of CQ resistance in *P. vivax* must be different.

P. vivax is not an exception. Rodent malaria parasites either naturally express (*P. yoelii*) or can be selected (*P. chabaudi*, *P. berghei*) for CQ resistance without the mutations in *pfcr*t or *pfmdr*1. The case of *P. chabaudi* is described in detail in section 1.7.2.

1.4.3. Molecular mechanisms of CQ resistance

Ever since the discovery of *pfcr*t it was speculated how exactly the mutations within that protein translate into the CQ-resistant phenotype. The involvement of a putative transporter, the reduced accumulation of drug by CQ-resistant parasite (Sanchez et al., 1997) and the reversal of resistance by calcium channel blocker (verapamil) (Martin et al., 1987) suggest that the altered CQ transport is the source of resistance. However, for a long time there was no consensus whether *pfcr*t transports chloroquine directly (either as an active efflux pump or as a channel allowing the passive leak out of vacuole) or indirectly (by modification of physiology of digestive vacuole). Today it seems that both of these may be true.

1.4.3.1. *Pfcr*t – a CQ transporter

The idea that *pfcr*t is a chloroquine transporter appeared as soon as it was discovered (Fidock et al., 2000). The series of experiments following the kinetics of labelled CQ accumulation in sensitive and resistant parasites indicated the presence of energy-dependent, saturable CQ efflux in resistant parasites that was dependent on the *PFCRT* K76T mutation (Sanchez et al., 2004; Sanchez et al., 2005; Sanchez et al., 2007a). Moreover both mutated and wild-type *pfcr*t forms expressed in yeast specifically bound CQ (Zhang et al., 2002; Zhang et al., 2004). Finally very recently it was demonstrated that *PFCRT* 76T (but not *PFCRT* 76K) protein introduced into the *Xenopus laevis* system can transport CQ across the oocyte membrane, supplying the direct evidence that altered CQ transport by *pfcr*t is a major factor contributing to the resistance (Martin et al., 2009) (Fig 1.8a).

Although initially it was suggested that *pfcr*t is an active transporter using energy to pump the drug from vacuole (Naude et al., 2005; Sanchez et al., 2007b) currently the accumulating body of evidence suggests that it is rather a gated channel allowing the passive leak of the protonated CQ along the concentration gradient (Bray et al., 2006; Cabrera et al., 2009a; Paguio et al., 2009) (Fig 1.8b).

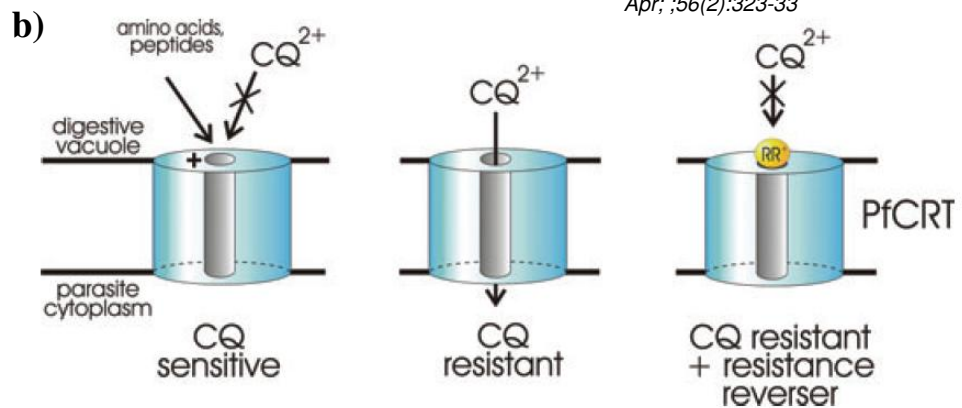
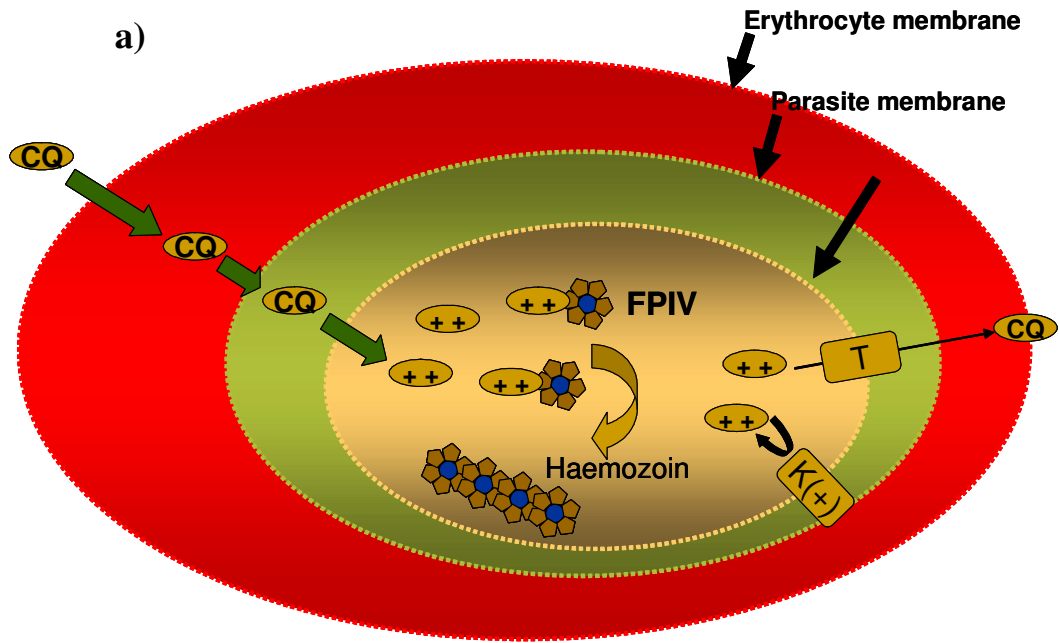
What is the relationship between *pfcr*t mutations and chloroquine transport? The analysis of *pfcr*t structure indicates that the crucial K76T mutation is in the part of the protein responsible for substrate recognition in the potential channel (Martin and Kirk, 2004). It is therefore possible that a wild-type version of *pfcr*t would

remain impermeable for the drug mainly because of the basic K76 residue situated at the entrance of the channel having a refractory role towards also positively charged CQ forms (Warhurst, 2003). The replacement of K76 with T would remove this obstacle and chloroquine would freely leak from the vacuole. It would also explain the role of verapamil (a channel blocker) in the resistance reversal. It would inhibit the function of this carrier, thus restoring its impermeability to CQ and retaining the high concentration of drug in the vacuole. Recently the transporter hypothesis was further supported by the fact that the introduction of an alternative basic residue into the entrance of the channel (S163R substitution) is also reversing the effect of K76T mutation (Johnson et al., 2004).

1.4.3.2. *PFCRT* - an alternative mechanism of resistance?

The transporter role of *pfcr*t is certainly one of the main components of CQ resistance. However it seems that a supplementary mechanism may be involved. It has been suggested that the reduced CQ accumulation cannot explain all of the CQ resistance – even with the similar intracellular CQ concentration the resistant parasites are still less sensitive to CQ than the sensitive ones (Cabrera et al., 2009b). Moreover the mutant *pfcr*t seems to alter the physiology of a digestive vacuole (Gligorijevic et al., 2006) and increase its acidity (Bennett et al., 2004). It was speculated that these changes can also have an impact on CQ resistance, by altering the bioavailability of its substrate (Roepe, 2009).

Finally, the studies of the reaction of resistant and sensitive parasites on the bolus CQ dose served at the different stages of synchronised culture did show that CQ-resistant parasites have *pfcr*t-related, decreased sensitivity to chloroquine not only in the trophozoite stage but also as a schizont or rings (that don't possess a fully formed digestive vacuole) (Gligorijevic et al., 2008). How *pfcr*t confers the CQ resistance to these stages remains unknown.



Bray et al. Mol microbiol. 2005
Apr; ;56(2):323-33

Figure 1.8 The possible role of PFCRT in CQ resistance

The protein interferes with CQ action by allowing the leak of charged drug out of the vacuole (a). It is caused by replacement of positively charged residue by the entrance of the channel with a neutral one (b) allowing the CQ transport. The calcium channel blocker reverses this process by blocking the transporter function of *pfcr*.

1.5. Artemisinin resistance

1.5.1. The resistance to artemisinins – an incoming threat

Soon after artemisinins were introduced, the general optimistic approach was that resistance to these drugs might not evolve. Its short half-life (most of the derivatives are metabolised within ~ 2 h (Meshnick et al., 1996b)) gives parasite very little time of exposure on suboptimal concentration of drug, which is necessary for development of the resistance (Stepniewska and White, 2008). Moreover artemisinins are generally used in combination with the other partner drugs which, in theory, should delay the resistance as described in section 3.2.1. Partly as a result of this optimistic prediction, the pipeline of development of new antimalarials was somewhat neglected and current drugs portfolio relies heavily on different versions of ACT (Olliaro and Wells, 2009). Although three different artemisinin derivatives are used, they are structurally similar and a significant level of cross-resistance is to be expected between them. Therefore we are currently relying on one drug and its loss would deal a major blow to malaria control programs.

This is particularly concerning in view of the recent reports of ART resistance selected *in vitro* and observed in the field, suggesting that the development of full-blown resistance is likely to be a matter of time.

1.5.2. The emerging artesunate resistance in central Asia

A few years ago initial reports of the reduced effectiveness of ACTs in western Cambodia started to accumulate (Denis et al., 2006; Alker et al., 2007), but their interpretation was complicated by the degree to which these data reflected pre-existing resistance to non-artemisinin components such as mefloquine. More recently however, well-controlled studies have been performed, monitoring *in vivo* and *in vitro* responses to artesunate and reporting treatment failures and reduced parasite clearance rates in patients with adequate plasma drug concentration (Dondorp et al., 2009; Noedl et al., 2009) thus confirming the presence of resistant strains in the population. Although currently the level of resistance seems to be low

and limited to a small region, even with proposed countermeasures (Dondorp et al., 2010) it is very likely to spread and increase its levels.

Moreover after years of unsuccessful attempts, an artemisinin resistant *P.falciparum* strain was generated under laboratory conditions (Chavchich et al., 2010). Three strains of slightly different resistance profiles were selected confirming the possibility of the rise and spread of ART resistance.

1.5.3. Genetics of ART resistance

Because ART-resistant strains were observed only recently, our knowledge of the genetics of this phenomenon is limited. The initial investigations performed using field isolates from the Thai-Cambodian border, revealed a high heritability of ART responses, suggesting the existence a strong genetic component of this trait (Anderson et al., 2010). However, there are no generally agreed resistance markers although a number of gene candidates have been proposed.

1.5.3.1. *Pfmdr1*

The studies of the influence of previously known resistance genes on the response to artemisinins have shown that the amplification of *pfmdr1* gene is connected with decreased ART sensitivity *in vitro* (Sidhu et al., 2006) and with the failure of artesunate-mefloquine combination *in vivo* (Alker et al., 2007) (although the latter could be due to the mefloquine component (Price et al., 2004)). The increase of *pfmdr1* copy number was observed also in two out of three ART resistant strains selected in laboratory conditions (Chavchich et al., 2010). The analysis of resistant strains isolated from the Thai Cambodia border however seem to exclude this gene as the main determinant of resistance in that setting (Dondorp et al., 2009; Imwong et al., 2010). In summary, like in the case of chloroquine, *pfmdr1* seems to influence artemisinin responses to some extent; its action however, is likely to be highly dependent on the other genetic factors.

1.5.3.2. *PfATP6*

PfATP6 (sarco/endoplasmic reticulum membrane calcium ATP-ase) was proposed as another candidate, as it was discovered to be a main target of ART action in *Xenopus* oocysts (Eckstein-Ludwig et al., 2003) and the introduction of a

single L263E mutation within this gene is sufficient to obtain resistance in this model system (Uhlemann et al., 2005). Additionally an association between a polymorphism in this gene (S769N) and the artemether responses *in vitro* was also suggested in isolates from French Guiana.(Jambou et al., 2005) However again no obvious association between this gene and ART responses was found in Southeast Asia (Dondorp et al., 2009; Imwong et al., 2010) and recent allelic exchange experiment introducing the *PfATP6* L263E mutation into various sensitive parasites failed to produce a significant change in ART responses. Therefore, *PfATP6* can be rejected as universal resistance determinant (although it cannot be excluded that it does influence the ART responses on some genetic backgrounds).

1.5.3.3.ubp1

A genetic cross between ART –resistant and sensitive *P.chabaudi* parasites indicated the possible involvement of locus on chromosome 2 and a gene encoding a deubiquitinating enzyme was found to be mutated in between the resistant and sensitive parasite (Hunt et al., 2007). The causative association between this gene and ART resistance was however uncertain, as the said mutations did appear before the exposure to artemisinin and no further genes within the implicated region were investigated so far. Additionally no mutations within this gene were found in any field or laboratory selected artemisinin-resistant *P.falciparum* parasites (Chavchich et al., 2010; Imwong et al., 2010).

1.6. Identification of genetic markers of resistance

1.6.1. Why are we interested in resistance markers?

Genetic markers of drug resistance are valuable tools for malaria control. From a practical point of view they allow a quick and easy assessment of the resistant status of laboratory and field isolates (see section 1.3.1.3). This information can be used for both more effective case management and for continuous monitoring of levels of resistance in the field. Moreover their identity can provide useful information concerning the mechanism of drug action, parasite biology, resistance evolution etc. that may lead to improved treatment policies or new drug discovery avenues etc. Therefore the rapid identification of resistance markers is of utmost importance for malaria research.

1.6.2. Methods of marker identification.

1.6.2.1. Candidate gene approaches.

In some cases, there may be pre-existing evidence suggesting the involvement of specific candidate genes or pathways. In such cases, the influence of these genes may be validated both in the field and in the laboratory. For example molecular targets of antifolates were known from studies of bacteria in 1930's and -40's (Hyde et al., 2008). Therefore the genes encoding these proteins were natural candidates and their investigation led to a discovery of the associations between mutations in *dhfr* and *dhps* genes and antifolates resistance

Today in the era of advanced molecular techniques (e.g. transfection) the verification of the influence of a specific gene on a resistance phenotype is relatively easily confirmed (e.g. involvement of *pfmdr1* in CQ resistance (Foote et al., 1990)) or rejected (like in case of ART resistance candidates (Imwong et al., 2010)). This method is however heavily dependent on pre-existing knowledge of gene candidates. In practice it is limited to drugs of either known mechanisms of action (like antifolates) or the similarities to other already studied resistance system (like *pfmdr1*). Moreover it concentrates on one particular gene at the time, which in the

case of multigenic traits may be very limiting. In cases such as these, other methods of mapping the genes conferring the resistance phenotypes are required, as described below.

1.6.2.2. Experimental linkage analysis

Classical genetic linkage analysis characterises the phenotype and genotype of individually cloned recombinant progeny of a genetic cross between genetically different resistant and sensitive parasite clones, and seeks genetic markers which are consistently inherited along with the resistance phenotype.

Three crosses of *P.falciparum* have been performed (Walliker et al., 1987; Wellems et al., 1991; Hayton et al., 2008). Each of them has provided valuable data concerning the genetic basis of drug resistance. For example, the first was used to confirm that the pyrimethamine resistance is caused by the mutations in *dhfr* gene (Peterson et al., 1988) and that *mdr1* modulates responses to many drugs (Duraisingh and Cowman, 2005). The second identified *pfcr* as the determinant of CQ resistance (Fidock et al., 2000). Finally the third is being used to investigate the genetic basis of quinine resistance.

Linkage analysis does not require any prior knowledge of the mechanism of resistance and the fact that all the progeny are carrying the combinations of just two genotypes allows us to eliminate the excessive genetic noise observed with genetically diverse isolates from the field (Hayton and Su, 2008). However in the case of *P. falciparum* technical and ethical constraints are making the generation of the crosses difficult. Another issue is that clearly characterised resistant and sensitive strains are needed for this analysis and sometimes (as is the case of artemisinin) stable genetic resistance may be difficult to obtain under laboratory conditions (Nzila and Mwai, 2009). Moreover, the results are very variable, largely depending on the number of isolated recombinants, quality of the cross, the frequency of genetic recombination in proximity of resistance locus etc.

1.6.2.3. Mapping determinants of drug resistance in field isolates

In some circumstances, the genetic determinants of drug resistance may leave a selection signature in the genomes of populations of parasites. For instance, where a particular (resistance) genotype spreads through a population under the influence of

drug use, genetic diversity will be reduced in resistant parasite populations relative to sensitive parasites (Anderson, 2004). Although genetic recombination may progressively introduce genetic diversity into the resistant population, there will be persistent reductions of diversity and areas of linkage disequilibrium in genes linked to those conferring the resistance phenotype. Such 'selective sweeps' have been observed for CQ and *crt* (Wootton et al., 2002) *dhfr* and antifolates (Nair et al., 2003) and *pfmdr1* (Nair et al., 2007) for mefloquine. Recently the combination of the above and the development of large-scale genotyping technologies resulted in the application of malaria genome-wide association analysis (GWAS) that resulted in mapping loci under selection with different drugs (Mu et al., 2010).

The main advantage of these studies is that they propose candidate markers with likely diagnostic value in the field without any prior knowledge regarding the genetics or mechanism of action of the drug. Moreover they take into account the entire genetic spectrum of the parasite population allowing the identification of many alleles and relationship between them while supplying data about the evolution and spread of resistance alleles.

The major drawback of this strategy is that it can be used only when the wide array of drug resistant strains are already present in the population (retrospective analysis), which means that the resistance is well established in the wild. It would be preferable to identify the genetic markers before (predictive analysis), to enable early intervention and control.

1.7. The *P.chabaudi* model of CQ and ART resistance

1.7.1. Why use the rodent model?

As rodent malarias share many similarities with human parasites in terms of life cycle, physiology and structure (Carter, 1978) they are often used as a convenient laboratory model. So far four African species of *Plasmodium* (*P.chabaudi*, *P.berghei*, *P.yoelii* and *P.vinckei*) have been adapted to growth in laboratory mice. They have supplied valuable data concerning the parasite life cycle (Talman et al., 2004; Poudel et al., 2008), host immune responses (Li et al., 2001; Cheesman et al., 2010a), infection dynamics (Bell et al., 2006; Reece et al., 2008) and pathogenesis (de Souza and Riley, 2002; Lamikanra et al., 2007).

The murine malaria model has some critical advantages. Principally, it provides a tractable experimental system to study the *Plasmodium* infection *in vivo*. The parasites can grow in their mammalian host in which they are exposed to the range of factors that are important for infection development but difficult to replicate in the culture flask (like anaemia, action of immune system, shearing forces caused by movement of parasites in the blood vessels, etc.). Moreover the rodent models give the possibility to safely follow the whole parasite lifecycle (including mosquito transmission) in laboratory conditions without the ethical constraints surrounding the use of the primates to study human malaria. Finally the recent publication of three rodent malaria genomes (*P.chabaudi*, *P.berghei* and *P.yoelii*) (<http://plasmodb.org/plasmo/>) revealed high degrees of synteny and gene conservation between rodent and human parasites (Carlton et al., 2002; Hall et al., 2005; Kooij et al., 2005) and provided a tool facilitating genetic and genomic investigations using these parasites.

In the case of the genetics of drug resistance, drug-resistant rodent malaria parasites are conveniently selected by the passage of parasites with increasing drug pressure (Carlton et al., 2001). Moreover, as both parental strains and selected mutants can be analysed, the relationship between the appearance of mutations and drug-resistance phenotypes may be studied within the same genetic background. In

addition the influence of host genetic factors can be avoided by using inbred mice strains. Finally, pharmacokinetics and pharmacodynamics processes in rodent and human host may be similar, even though their kinetic parameters differ (Cambie et al., 1994; Batty et al., 2008).

Although all rodent malarias have been used to investigate the action of, and resistance to, drugs (e.g. *P.berghei* (van Dijk et al., 1994; Gervais et al., 1999; Syafruddin et al., 1999), *P.yoelii* (Cheng and Saul, 1994; Srivastava et al., 1999) and *P.vinckei* (Powers et al., 1969; Puri and Chandra, 2006)), *P.chabaudi* has been most exploited for the genetic analysis of genetically stable phenotypes (Carlton et al., 2001).

1.7.2. The lineage of resistant *P.chabaudi*

The work presented here exploits an ‘isogenic’ lineage of drug resistant *P.chabaudi*, selected over the years in David Walliker’s laboratory for resistance to various antimalarials (Fig. 1.9). The original all-sensitive **AS-sens** strain is a cloned isolate obtained in 1969 from Thicket rat (*Thamnomys rutilans*) in Central Africa and adapted to growth in laboratory mice (Carter and Walliker, 1975). From that strain with single step pyrimethamine treatment (*i.e.* one mouse passage), the pyrimethamine resistant strain **AS-pyr** was selected (Walliker et al., 1975). This strain was subsequently used for two different selections. On one hand a single step selection for increased resistance to pyrimethamine/ sulfadoxine combination resulted in AS-50SP strain (Hayton et al., 2002). On the other hand, the parasites were exposed to increasing dose of chloroquine over five consecutive mouse passages, generating the **AS-3CQ** strain (Rosario, 1976). The additional selection of the latter with higher chloroquine doses resulted in **AS-15CQ** and **AS-30CQ** – an intermediate and highly resistant strain selected over 10 and 35 passages respectively (Padua, 1981). The AS-15CQ strain was further selected for mefloquine resistance for 23 mouse passages (Cravo et al., 2003) generating **AS-15MF** strain. Later, AS-15CQ and AS-30CQ were selected for increased artesunate and artemisinin resistance respectively for 14 mouse passages (Afonso et al., 2006) resulting in **AS-ATN** and **AS-ART** strains. All strains, except **AS-15CQ** were cloned by limiting

dilution. In each case the resistance was stable through >20 passages in the absence of drugs, mosquito transmission and deep-freezing.

Interestingly, during the process of selection for artemisinin resistance it was revealed that the AS-15CQ and AS-30CQ already possessed an increased level of tolerance to this drug (artemisinin resistance phenotype I) in comparison to the AS-3CQ and AS-sens that were sensitive. Further selection resulted in artemisinin resistance phenotype II observed in **AS-ART** (unpublished observations). Therefore in the *P.chabaudi* lineage the resistance to artemisinin arose spontaneously under chloroquine selection without previous contact with this drug.

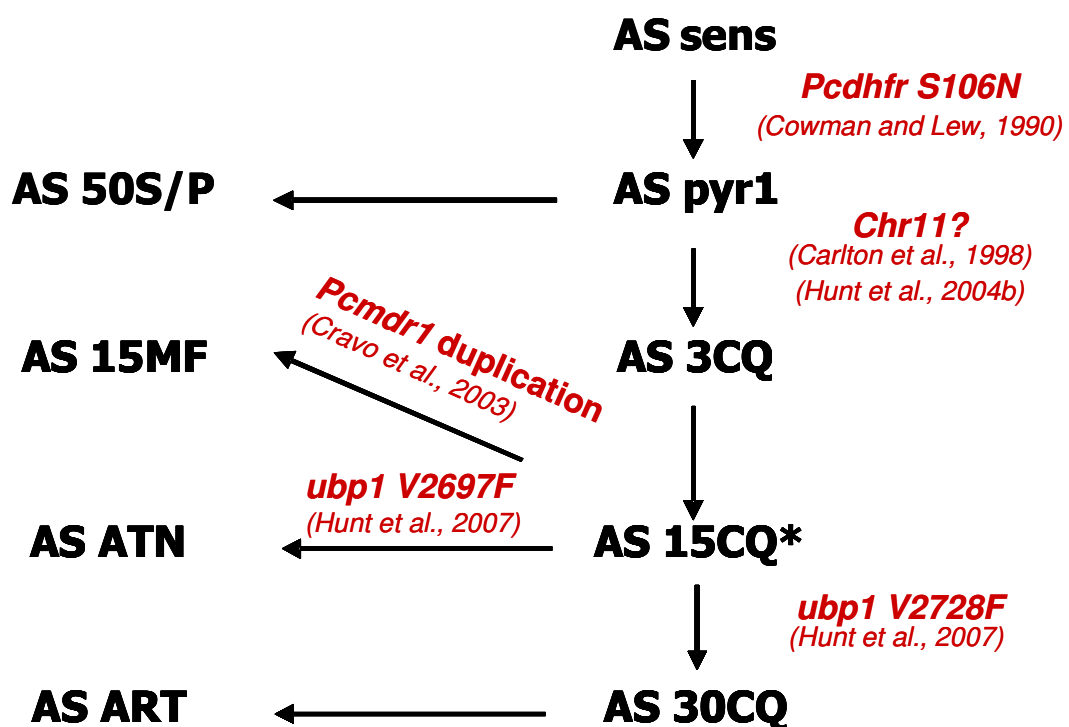


Figure 1.9 The resistant AS lineage of *P.chabaudi*

Each arrow represents the selection step, the known mutations/loci with potential role in the resistance are marked on red, *- uncloned strain.

1.7.3. Past investigations of the genetics of drug resistance in *P.chabaudi* lineage

The AS lineage of parasites (Fig 1.9) has been used to investigate the genetic basis of resistance to pyrimethamine, S/P, CQ, MF and ART, and to map the underlying genes generally by the application of classical linkage analysis: recombinant clones of genetic crosses are analysed by *in vivo* phenotyping of drug response, and genotyping in order to determine the genetic markers inherited along with drug response.

1.7.3.1. Pyrimethamine resistance

The candidate gene approach was used to study the genetic background of resistance leading to a discovery of S106N mutation *dhfr* gene (Cowman and Lew, 1990), situated on chromosome 7. Later an advanced linkage analysis confirmed that the *dhfr* locus is the main determinant of resistance (Culleton et al., 2005).

The identified gene is an orthologue of *pf dhfr* which has previously been connected to pyrimethamine resistance in *P.falciparum* (Peterson et al., 1988). Also the causative mutation S106N is analogous to the S108N conferring resistance in human malaria.

1.7.3.2. Chloroquine resistance

Chloroquine resistance was selected by multiple passages in increasing drug concentration and three strains of different level of resistance were obtained (AS-3CQ, AS-15CQ and AS-30CQ). The cross performed between the low resistant AS-3CQ and the genetically different, sensitive strain (AJ) established that resistance had genetic basis. Moreover, the typical Mendelian inheritance of this trait suggested the existence of a single locus responsible for low CQ resistance (Rosario, 1976). Linkage analysis of AS-3CQxAJ cross indicated that the location of this main locus of resistance was on chromosome 11 (Carlton et al., 1998). Later, with addition of more genetic markers, this region was narrowed to ~ 250 kb fragment containing ~50 different genes (Hunt et al., 2004b) but no obvious gene candidates were proposed.

A genetic cross was also performed between the highly resistant AS-30CQ strain and AJ. This time intermediate levels of resistance were observed in recombinants suggesting that the high CQ resistance is multigenic. No additional loci have been identified.

The two genes responsible for chloroquine resistance in *P.falciparum* – *pfcr1* and *pfmdr1* were investigated for their role in chloroquine resistance in *P.chabaudi*. Firstly, no mutations in either gene were found in the AS lineage (Hunt et al., 2004a). Secondly, there was no linkage in genetic crosses (AS-3CQ x AJ or AS-30CQ x AJ). Therefore the involvement of these genes in CQ resistance in *P.chabaudi* is highly unlikely.

Interestingly, despite being founded of a different genetic basis, CQ resistance in *P.chabaudi* seems to share a lot of features with those observed in human malaria including the reduced CQ accumulation in resistant parasites (Miki et al., 1992), changes in the morphology of digestive vacuoles (Ohsawa et al., 1991) and reversal of resistance by calcium channel blockers like verapamil (Tanabe et al., 1990). The orthologues of causative *P.chabaudi* genes may prove to be additional factors influencing CQ response in *P.falciparum* or the main determinants of resistance in *P.vivax* (in which the genetic basis of resistance is still unknown and unlinked to *pfcr1*).

1.7.3.3.Mefloquine resistance

In *P.falciparum* mefloquine resistance was associated with mutations (Reed et al., 2000) and amplification (Cowman et al., 1994; Peel et al., 1994) of *pfmdr1* gene. In *P.chabaudi*, the mefloquine-resistant clone AS-15MF also bears a duplication of the *mdr1* gene, and this event was linked to mefloquine resistance in the progeny of the AS-15MF x AJ cross although other genes may be involved (Cravo et al., 2003).

1.7.3.4.Artemisinin resistance

In the case of artemisinin resistance, the data from the rodent model cannot be compared to human parasites because the determinants of artemisinin responses are not known yet. However previous work on *P.chabaudi* indicated the existence of one main locus of artemisinin resistance is situated on chromosome 2 (Hunt et al., 2007). Within this locus the *ubp1* gene was discovered. It was mutated in both AS-

ATN (V2697F mutation) and AS-ART (V2728F mutation) parasites. However, both mutations did appear during the CQ selection in AS-15CQ parasite and, its direct role in the resistance phenotype is still not fully understood. For instance, the possibility that there are other mutations in genes linked to *ubp1* was not explored.

1.8. New gene identification strategy – a combination of Linkage Group Selection and next generation sequencing on isogenic drug resistant lines of parasite

The aim of this section is to describe the approach that I have taken to identify novel genes involved CQ and ART resistance. It uses a variant of genetic linkage analysis called Linkage Group Selection (LGS), in order to map the potential loci of resistance and next generation sequencing of isogenic drug resistant parasites to identify specific mutations within these loci. It is anticipated that the genes identified can be evaluated in *P.vivax* and *P.falciparum* in the field and in laboratory transfection studies.

1.8.1. Linkage Group Selection – a fast method to map resistance loci

In contrast to classical genetic linkage analysis, which maps genes by defining the co-inheritance of genetic markers and the phenotype in *individually* cloned recombinant progeny, Linkage Group Selection (Carter et al., 2007) (LGS) analyses the ‘phenotype’ and ‘genotype’ of *large populations* of uncloned recombinant progeny. It applies a selection (such as drug treatment or host immunity) to the pooled progeny of a genetic cross and analyses the proportions of parental alleles at many polymorphic loci in the selected and unselected populations.

LGS was previously developed by Richard Carter and collaborators and applied to pyrimethamine resistance and strain-specific immunity in *P. chabaudi* (Culleton et al., 2005; Martinelli et al., 2005). It has subsequently been successfully applied to artemisinin resistance (Hunt et al., 2007), strain specific immunity in *P.chabaudi* (Cheesman et al., 2010b) and growth rate in *P.yoelii* (Pattaradilokrat et al., 2009). Another approach called Bulk Segregation Analysis has some similar features and was also used for the successful identification of genes underlying various traits in yeast (Wenger et al., 2010) and maize (Liu et al., 2010)

The basic elements of the LGS strategy are discussed in the three sections below and presented in Fig .1.10

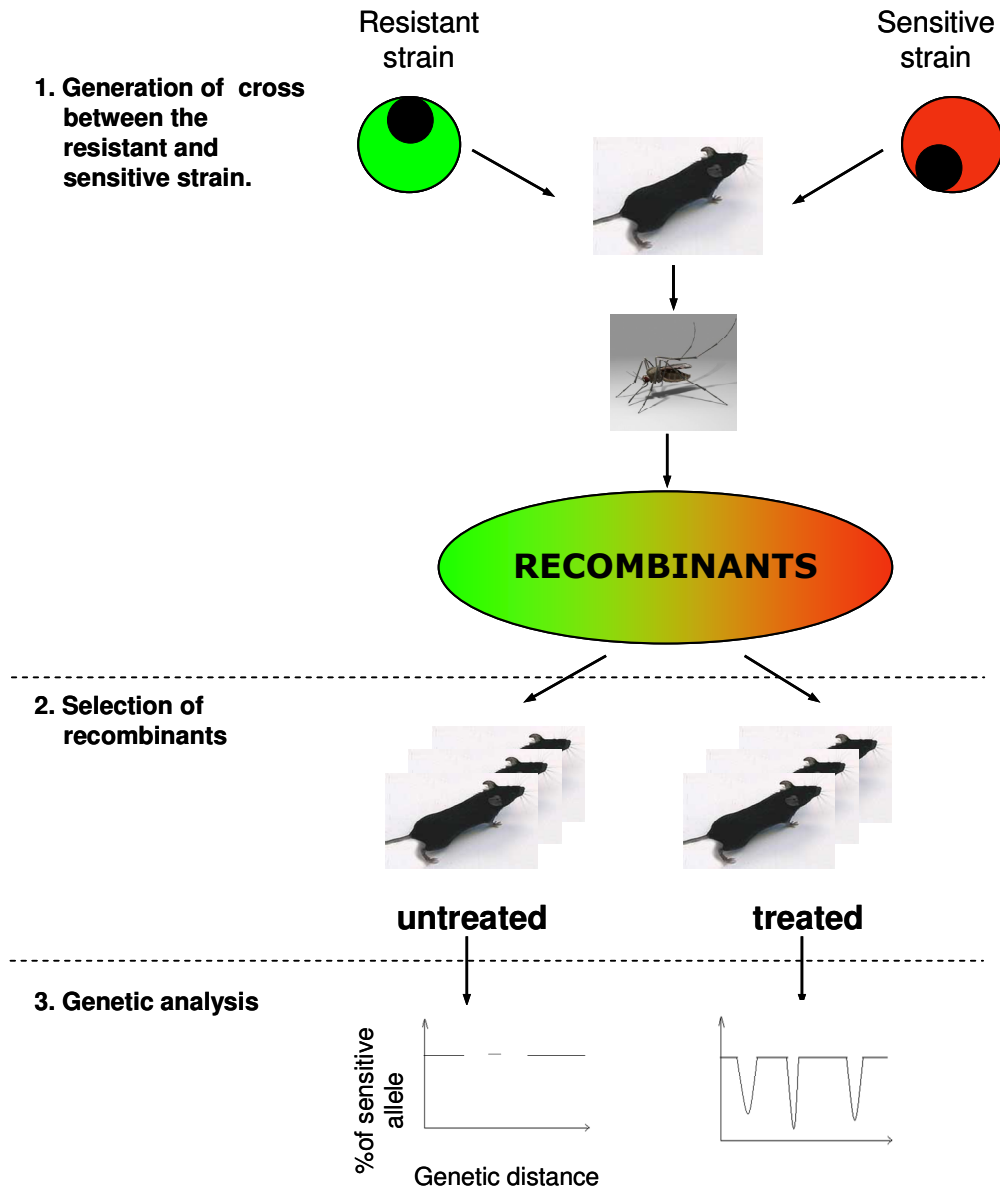


Figure 1.10 The design of a Linkage Group Selection experiment
 Explanations in the text.

1.8.1.1.Generation of experimental cross

As in classical linkage analysis, LGS requires a genetic cross between two unrelated parasites, one of sensitive and the other resistant to the relevant selection pressure (*e.g.* drug treatment). In *P. chabaudi* the cross is generated by infecting the mice with a mixture of resistant and sensitive strains of the parasites and feeding them to *Anopheles* mosquitoes. In a mosquito's midgut the gametes from both strains fuse to form a zygote. Recombination results in the formation of new parasites (parentals and recombinants). These can be used to infect new mice and to recover progeny as haploid asexual parasite forms in mouse red blood cells.

1.8.1.2.Selection

In contrast to the classical linkage analysis, where independently cloned parasites are phenotyped (*e.g.* tested for drug response), in LGS, a population approach is used. The pooled, uncloned progeny are grown in two groups of mice, one treated with drug and the other untreated.

If the cross was successful initially both groups will contain the parasites populations containing mixture of parental alleles at all loci according to the principles of Mendelian genetics. Later however the structure of both populations will change reflecting the relative fitness of various genotypes in the growth environment. Assuming that the drug resistance was only phenotyping difference between the parental strains, in the untreated group most parasites will have equal chances to grow and only minimal deviation from initial population structure will be observed.(Fig 1.11a). In treated animals however, the drug-sensitive recombinants and the sensitive parental parasite will be eliminated by drug treatment (Fig 1.11b). It generates a population in which all the parasites will carry the alleles of resistant parent in fragments of the genome linked to the resistance.

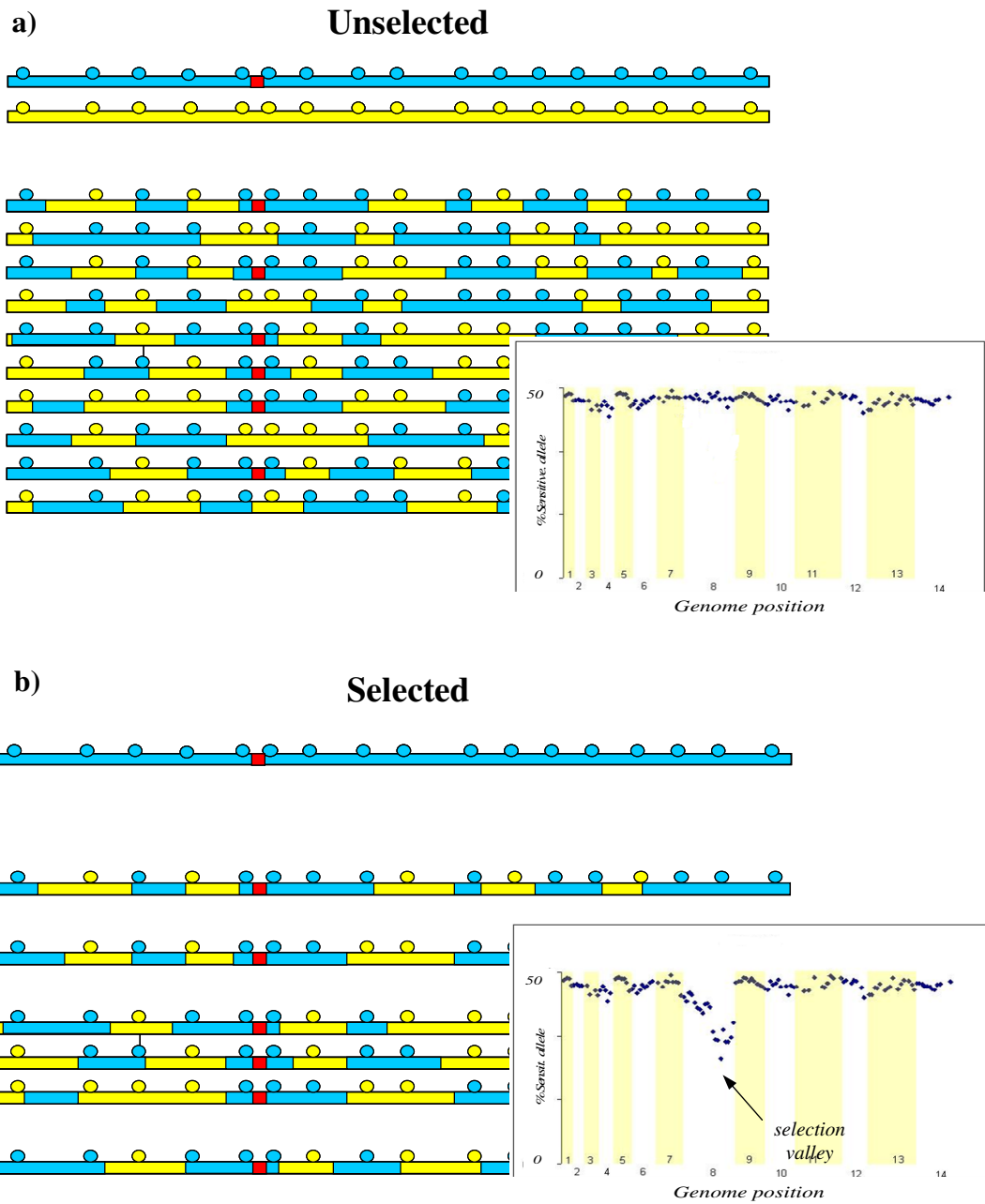


Figure 1.11 The possible results of genetic analysis

The unselected (a) and selected (b) parasite population and their genome scans. The bars are representing fragments of sensitive (yellow) and resistant (blue) genotypes. Circles are representing the genetic markers. The scans are results of plotting the frequency of markers from resistant parent against the genome position

1.8.1.3. Genetic analysis

The frequency of genome-wide genetic markers from the sensitive parent is studied in the DNA from the pooled progeny and plotted along the genomic distance presenting a so-called “genome scan”.

In the unselected population, the proportion of markers across the genome is expected to reflect the initial proportions of the two parentals (typically 50%). Therefore a proportion of markers of each parent plotted against the genomic distance will be represented by a straight line (Fig 1.11a). In the selected population however, the markers from the sensitive parent that are linked to the locus of drug resistance tend to be eliminated from the population and this effect is related to the genetic distance between the marker and the resistance locus. In this situation we will observe a “selection valley” indicating the causative locus (Fig 1.11b).

While classical genetic linkage analysis employed any ‘binary’ genetic or protein marker, such as enzyme mobility (Rosario, 1976), RFLP (Carlton et al., 1998), AFLP (Grech et al., 2002), microsatellite or sequenced SNP; for LGS, quantitative markers are required to estimate allele proportion in selected population. Initially AFLP markers were adapted to LGS analysis (Martinelli et al., 2004). Now however, to avoid some disadvantages and inconvenience of these assays, quantitative SNP PyrosequencingTM is used (Cheesman et al., 2007) (see section 3.1.3).

1.8.2. Next generation sequencing – a method of mutation discovery

The existence of each selection valley indicates the presence of a ‘major effect’ gene conferring resistance to the selection pressure applied in LGS. Because the bottom of a selection valley may extend over about 100 kb, large numbers of gene candidates are generated by the selection scan. Previously, candidate gene approaches were used to search for the resistance gene and a significant amount of sequencing work was required in order to specify the causative mutation. This approach however required a lot of effort and (unless a single obvious candidate could be localised in the selection valley, based on the knowledge of biology of the

studied phenomenon or analogy with other species), usually did not lead to discovery of causative mutation.

Now for the first time, high-throughput genome re-sequencing was employed in order map the causative mutations within the selection valleys. The Illumina Solexa genome analyser is a sequencing system that allows rapid and accurate sequencing on a genome scale. It is based on the short reads (>50bp) that can be sequenced and mapped to the pre-existing genome scaffold (here the published *P.chabaudi* genome). Differences between the ‘assembled’ sequence and the ‘reference’ sequence can be used to identify candidate mutations.

In this work Solexa is used for identification of genetic changes between two congenic strains in order to identify the genes mutated within predicted loci. Both resistant mutant and its sensitive progenitor can be sequenced and all the genetic changes that occurred during the drug selection in the locus indicated by LGS can be identified.

In summary the Linkage Group Selection indicates the regions of the genome linked to the resistance and Solexa genome resequencing provides the list of mutations that appeared in the strain under drug pressure. If there is only one change that appeared with a selected locus under drug pressure, it is a very strong candidate for the resistance-conferring mutation, although additional studies (*e.g.* transfection of the mutated gene-candidate into drug-sensitive parasite) are required to unambiguously prove this connection.

1.9. Aims and content of this thesis

This thesis describes how Linkage Group Selection and Solexa genome re-sequencing were used to map the genes conferring different levels of CQ and ART resistance.

In Chapter 2, I describe the confirmation of chloroquine resistance phenotypes in the AS lineage and preliminary experiments that establish the optimum chloroquine concentrations for LGS experiments

In Chapter 3 and 4, I describe the generation of a cross and a subsequent backcross between the chloroquine and artemisinin resistant AS-30CQ strain and the genetically different, sensitive AJ strain, the selection of these crosses with previously established chloroquine and artemisinin concentrations and their genetic analysis. These experiments are designed to identify loci involved in increasing resistance to chloroquine and to confirm the existence of a locus responsible for resistance to artemisinin.

In Chapter 5, I describe the genome re-sequencing of AS-30CQ strain and its chloroquine and artemisinin sensitive predecessor – AS-sens strain in order to establish the full inventories of mutations accumulating between the two strains. This list, combined with LGS data from previous experiments allowed us to identify:

- A strong candidate for a main determinant of CQ resistance in *P. chabaudi*
- two possible candidates contributing to the intermediate CQ response
- the main determinant of ART resistance and the highest level of resistance to CQ

In Chapter 6, I investigate the fitness cost arising with high CQ resistance by comparing the growth of low and high-CQ resistant mutants in competition.

These data and their implications for understanding CQ- and ART resistance are discussed in Chapter 7.

2. The preliminary experiments – establishing the selection conditions for LGS experiments

2.1. Introduction

This chapter describes preliminary experiments performed in order to confirm the relative phenotypes of strains later used for the genetic cross and to establish the optimal selection conditions in LGS experiments.

2.1.1. The importance of the selection in LGS

In section 1.8 of Chapter 1, I proposed an approach to identify the loci of chloroquine and artemisinin selection in the rodent malaria parasite *P.chabaudi*. Its crucial component is Linkage Group Selection – the genetic analysis of the pooled progeny of a genetic cross between a resistant and a sensitive strain, after a drug selection.

The resistant parent of the cross described in chapter 3, AS-30CQ, has been derived from its predecessors of lower (AS-3CQ) and intermediate (AS-15CQ) resistance to CQ (Fig. 1.9, Chapter 1). Therefore, it is likely that at least three genes contribute, one after another, to the highest level of resistance (however other possibilities like the accumulation of epigenetic changes, or three completely different genetic mechanisms in each of the consecutive strains cannot be excluded). Additionally, the differences in growth rate (which is probably determined by many loci) exist between the two parents of the cross, AS-30CQ and AJ (de Roode et al., 2005). In the cross both parental genotypes will be reassorted during the meiotic recombination, producing parasites of different level of resistance and growth rate.

In order to select from that pool the parasites of a desired level of resistance, a carefully designed drugging regime (length of the treatment, drug dose, frequency) needs to be used. On the one hand, insufficient treatment would allow the growth of some sensitive recombinants which would interfere with further analysis. On the other hand, too much drug might kill some resistant parasites, not yielding enough recombinants in the pool of genotypes in the population. The length of infection may also be a factor because each day/round of sexual reproduction after the end of the

treatment when the parasites are already selected, the drug in serum is being metabolised and other factors (such as the mouse immune system, competition for remaining resources), can modify the structure of the population. The optimal drug treatment is likely, therefore, to be large enough for efficient selection, but low enough to yield large numbers of recombinant parasites in short time.

2.1.2. The requirements of different drugging regimes

My aim was to select the AS-30CQxAJ cross with different CQ concentration in order to map the loci required for different levels of resistance (this concept is fully explained in section 4.1.3 of Chapter 4). Therefore, I needed to establish at least two different drugging regimes. The first one would select for the highly resistant parasites (AS-30CQ phenotype) only, eliminating the sensitive and low resistant ones. This regime would have an advantage of providing a maximal level of selection on all the loci involved in resistance. However it is likely to provide less genetic material and, as the recrudescence of parasites is likely to take longer, be more susceptible to other selection pressures that may appear during infection (like the selection for faster growth or the efficient evasion of the immune system).

The second regime would select parasites exhibiting any level of resistance (even the low resistant AS-3CQ phenotype) and remove only the fully sensitive ones (like AJ phenotype). This would have the advantage of generating a larger parasite population containing more DNA and more diverse recombinants improving the quality of genetic analysis. However, the lack of complete selection on loci involved in the highest levels of resistance could make the latter one difficult to map.

Once these drugging regimes were established, the more suitable one or combination of both could be used for the selection of cross progeny.

2.2. Experimental design

The two parents of the cross: AJ (sensitive) and AS-30CQ (highly resistant) as well as AS-3CQ (the low-resistant progenitor of AS-30CQ) were used as models of sensitive, high-resistant and low-resistant progeny of the cross, respectively. The selection conditions were estimated using mixed infections of two strains treated with different drug concentration and monitored using a genetic marker between the two strains. At the time of this experiment no genetic marker distinguishing between AS-3CQ and AS-30CQ was defined. However, a proportional sequencing assay based on 4 SNP in *pcdhps* (PCHAS_142850) gene was available and could be used to measure the proportions between AJ and strains from the AS lineage in a single infection (Hunt et al, 2005).

Two competition experiments were therefore performed: AS-30CQ + AJ and AS-3CQ + AJ.

2.2.1. AS-30CQ+AJ

The aim of the first experiment (AS-30CQ +AJ) was to establish the best drugging regime to select the highly resistant recombinants. To this effect four different CQ concentrations were tested. The highest one was 30 mg of CQ kg⁻¹ day⁻¹ – the highly resistant strain AS-30CQ was selected using this concentration. Three lower concentrations (5,10 and 20 mg of CQ kg⁻¹ day⁻¹) were also employed – all of them higher than 3 mg of CQ kg⁻¹ day⁻¹ used to select the low resistant AS-3CQ strain. Additionally the influence of a different number of days of treatments was tested – 3, 4 and 5 day drugging regimes were used.

The plan of the experiment is presented on Fig 2.1: The 1:1 mixtures of highly resistant (AS 30CQ) and sensitive (AJ) clone were inoculated into 5 groups of mice, 3 mice each. Each group was drugged with a different CQ concentration – 0, 5, 10, 20 and 30 mg of CQ kg⁻¹ day⁻¹. The drugging process started on the day of infection (day 0 p.i.) and continued for respectively 3, 4 and 5 days for mice 1, 2 and 3 in each group.

Independently, single infection controls were prepared – 2 mice infected with AS-30CQ and 2 with AJ parasites – one of each group undrugged and one treated with 5 mg of CQ kg⁻¹ day⁻¹ for 4 days.

In summary only one mouse was used per each drug dose/treatment length/strain combination. This decision was taken, because the aim of these pilot experiments was to confirm the identity of the used strains and estimate the range of the concentrations that could be used for the actual selections, rather than quantify the parasites responses to particular treatment regimens (to this effect the experiments similar to the ones described in chapter 6 would have to be performed). Therefore, in the light of limited number of animals that could be managed during one experiment testing of various regimes even with one animal was judged more important than producing the biological replicates.

The parasitaemia and proportion of resistant and sensitive parasites in the infection were monitored for 16 days.

2.2.2. AS-3CQ+AJ

The first objective of the second experiment (AS3CQ+AJ) was to find the lowest drug concentration eliminating AJ and allowing the growth of low resistant parasites that could be used to separate all the resistant recombinants from sensitive ones. The second objective is to establish a concentration inhibitory for both the sensitive (AJ) and low-resistant strain (AS-3CQ) that would still allow the growth of the highly resistant one (AS-30CQ) (from previous experiment). Again four different concentrations were used:

- 1.5 mg of CQ kg⁻¹ day⁻¹ – the very low concentration, probably allowing the growth of all resistant strains
- 3 mg of CQ kg⁻¹ day⁻¹ – concentration used to select the low resistant AS-3CQ strain, therefore permissive for all resistant parasites
- 5 mg of CQ kg⁻¹ day⁻¹ – higher than the concentration used to select AS-3CQ but lower than 15 mg of CQ kg⁻¹ day⁻¹ used for selection of AS-15CQ and therefore

likely to select the intermediate and highly resistant parasites from the low resistant and sensitive ones

- 20 mg of CQ $\text{kg}^{-1} \text{ day}^{-1}$ – as inferred from selection data inhibitory for AS-15CQ and AS-3CQ but allowing the growth of AS-30CQ.

The design of the experiment is presented in Fig. 2.2. Five groups of 4 mice were infected with a mixture of AS-3CQ and AJ parasites, in proportion 50:50. All mice were drugged for three days with different chloroquine concentrations for each group (0, 1.5, 3, 5, 20 mg of CQ $\text{kg}^{-1} \text{ day}^{-1}$). In parallel, controls were set as described in the previous experiment: two mice were infected with AS-3CQ and two with AJ strain. One from the two was left untreated and one treated with 3 mg of CQ $\text{kg}^{-1} \text{ day}^{-1}$.

As the three day treatment regime was decided in the previous experiment and the range of the concentration to be tested was narrower this time mixed infections were performed in 4 replicates in order to assess the influence of the host factors on the selection.

The parasitaemia and proportion of both strains were monitored for 19 days.

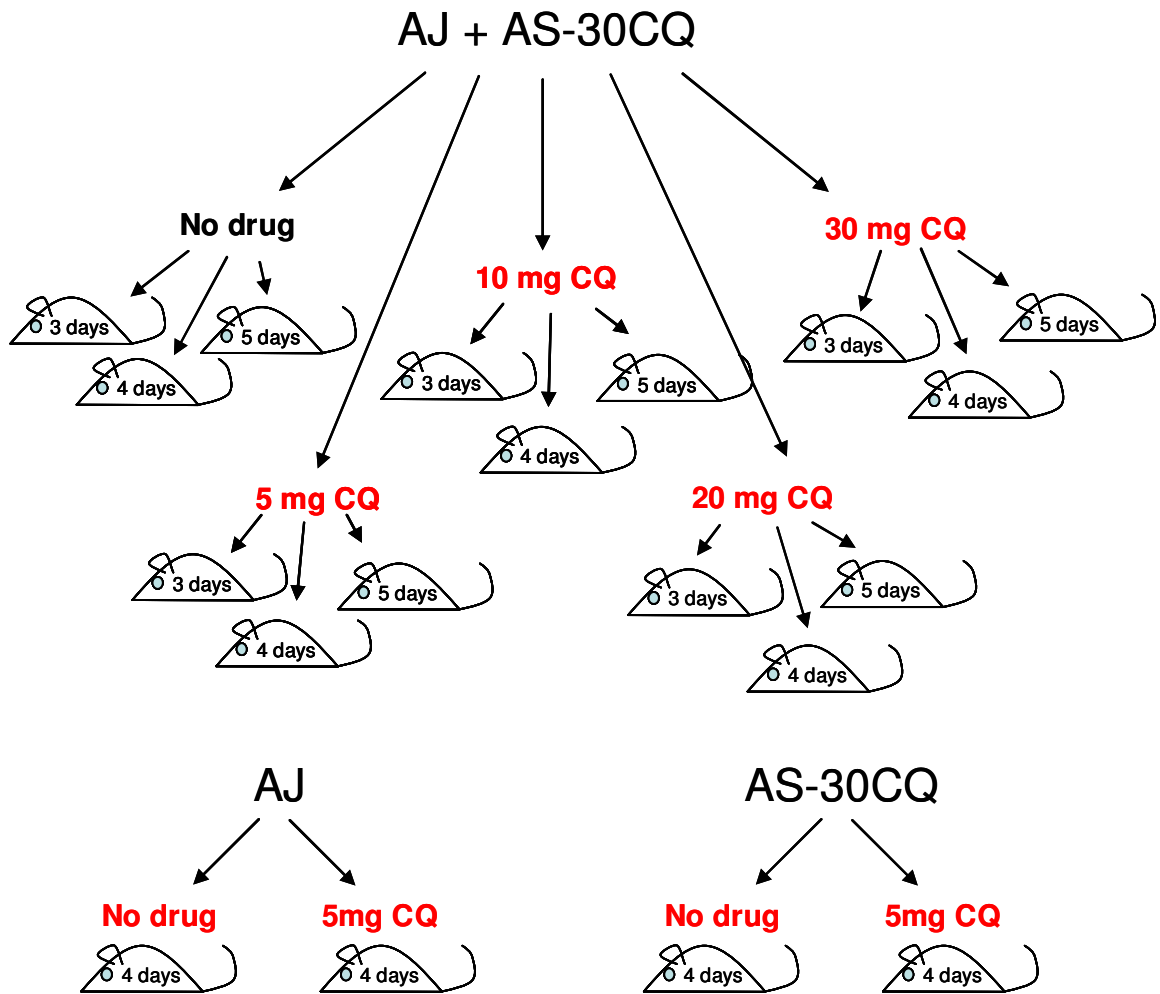


Figure 2.1 The design of AS-30CQ+ AJ experiment.

The numbers inside animals indicate the number of days of treatment. The description above – the concentration of drug used. Each mouse symbol represents one animal.

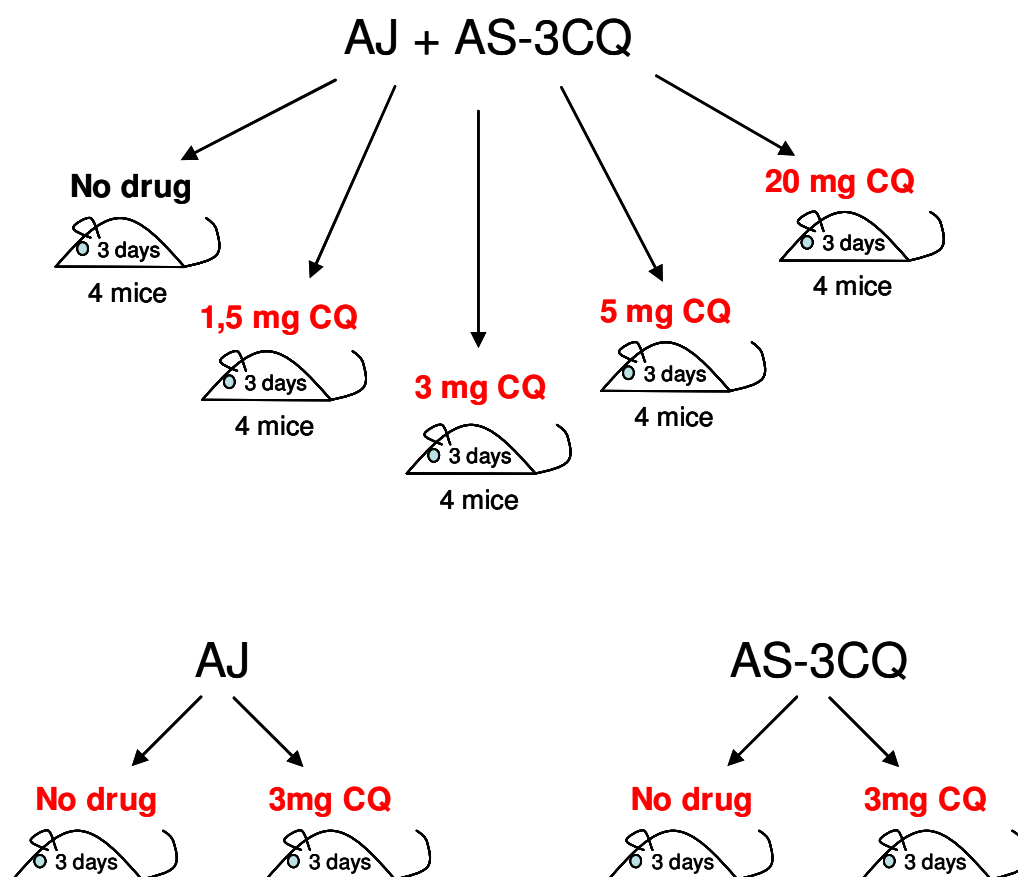


Figure 2.2. The design of AS-3CQ+ AJ experiment

The numbers inside animals indicate the number of days of treatment. The description above – the concentration of drug used.

2.3. Materials and methods

2.3.1. Parasite and mice strains

Three strains of *P.chabaudi* were used in this experiment. Two of them (the highly resistant AS-30CQ and low resistant AS-3CQ) are part of the resistant lineage selected from the original sensitive AS-sens strain, described in detail in section 1.7.2 of Chapter 1. Third one (AJ) is a genetically different, sensitive strain isolated in 1969 (Carter and Walliker, 1975) from central African thicket rat (*Thamnomys rutilans*).

All parasites were grown in 8-14 weeks old female CBA mice. The animals were housed at the University of Edinburgh animal unit (March Building) according to animal husbandry and veterinary standards. All the animal procedures were approved under the U.K. Home Office Animals (Scientific Procedures) Act 1986 (Project license number 60/3578, personal licence number 60/10747). The mice had unlimited access to food (RM3 diet, SDS UK Ltd.) and water enriched with 0.05% of para-aminobenzoic acid (PABA). All animals were observed and, if the symptoms of severe disease were noticed, the animal was killed according to the regulations of Schedule 1 humane killing.

2.3.2. Setting up the experimental infections

2.3.2.1.Preparation of inoculum

Initially parasites for inoculation were propagated from the cryopreserved samples. The capillary containing the aliquot of blood from infected animals, suspended in deep-freezing solution (see Appendix A), was defrosted in ice and the recovered blood was injected intraperitoneally. Once the parasites were observed in peripheral blood the mouse could be used as the donor for a larger group of experimental mice.

2.3.2.2.Single strain infections

On the day of infection the full cell blood count of a donor mouse was performed using a Coulter counter (Beckman Coulter, Inc.) and parasitaemia determined using the thin blood smear as described in section 2.3.4. This was used

to calculate the parasite density and, hence, the amount of blood containing the number of parasites required for infections (10^6 of parasites per 20 g mouse were used unless specified otherwise). The appropriate volume of blood was taken from a mouse tail vein using a thin glass capillary and diluted with citrate saline (see Appendix A) to contain the desired number of parasites in 100 μ l. This solution was intraperitoneally injected into new mice in a volume 100 μ l per 20 g of mouse bodyweight.

2.3.2.3. The mixed infections

Mixed infections were set as above except that both strains were propagated separately, diluted to 10^6 parasites / 100 μ l using the protocol from section 2.3.2.2. and mixed in a 1:1 proportion. Then the mice were injected with 100 μ l of this mixture/20 g of bodyweight

2.3.3. Drugging

The mice were treated with chloroquine sulphate (Beacon Pharmaceuticals), diluted to desired concentration with water and administered to the animals by gavage (100 μ l of solution/20 g of mouse body weight). The untreated animals were receiving an equivalent volume of water. All drugging started at day 0 and continued for 3, 4 or 5 days according to Fig. 2.1 and 2.2. The treatment took place the same time each day.

2.3.4. Monitoring the parasitaemia

Daily thin blood smears were taken from all the animals, fixed with methanol and stained with Giemsa's reagent (BDH) to monitor the parasitaemia.

Parasitaemia was determined by counting infected and total red blood cells on five representative microscope fields. At least 5 fields containing a minimum of 100 RBC each were counted. The parasitaemia was calculated according to the formula:

$$\text{Parasitaemia} = \text{Infected RBC} / \text{Total RBC}$$

The erythrocytes containing more than one parasite were counted as one infected cell.

2.3.5. Proportional sequencing of *pcdhps* gene

The proportions of the AS and AJ strains were established using the method of proportional sequencing. This method (described in detail in (Hunt et al., 2005)) estimates the proportions of SNPs in a mixture by measuring, on a DNA sequencing electropherogram, the heights of the peaks corresponding to two different bases of the polymorphism. In this case four bases within *pcdhps* gene, polymorphic between AJ and AS lineage were used. This assay was previously described and calibrated showing the good efficiency with experimental error of ~2% (Fig. 2.3 b) (calibration and efficiency of that particular assay are discussed in detail in Hunt et al., 2005).

2.3.5.1. Blood sampling and small scale DNA isolation

In order to monitor the proportion of the strain in mixed infections every second day 10µl of the blood from a mouse tail vein was taken on 180 µl of PBS (see Appendix 1) and frozen for genetic analysis. The animals were sampled every second day.

Later all the blood samples were defrosted and DNA was isolated using High Pure PCR Template Preparation Kit (Roche).

2.3.5.2. Proportional sequencing of *pcdhps* gene

The DNA from each sample was used as a template in triplicate PCR reaction amplifying the fragment of *pcdhps* gene containing four polymorphisms. The standard PCR protocol (see Appendix B) was used with following primers:

Outer PCR:

pcdhps-11 GTACGCAGAATATTTCAAATG

pcdhps-12 CCTTGAATACCCAATAAAAAG

Inner PCR

pcdhps-07 CTTTGTTCCTCATAATCCAG

pcdhps-08 GGTTTAGGTTTGCAAAGAA

The products were purified using QIAquick PCR purification kit (Qiagen) and sequenced (forward and reverse) using an ABI Prism0® BigDye™ Terminator dideoxy sequencing system (Applied Biosystems).

2.3.5.3. Estimation of proportions of parasites in the mixture

The sequencing results were analysed using Chromas 2.33 software (Technelysium Pty Ltd) and the heights of peaks corresponding to the two bases were measured on each of the four polymorphic positions using the “y” coordinate of the highest point of each peak (see Fig. 2.3 a).

The proportion of the AS strain (f_{AS}) was estimated for each of four SNPs (forward and reverse) with the following equation:

$$f_{AS} = k_p \times (h_{AS} \times 100 / (h_{AS} + h_{AJ}))$$

where h_{AS} and h_{AJ} are the heights of peaks corresponding to the AS and AJ alleles respectively for the further details and k_p is a coefficient calculated for each of the polymorphisms based on calibration data, reflecting the systematic error of the measurement (tendency to over- or underestimate the proportions of the two strains).

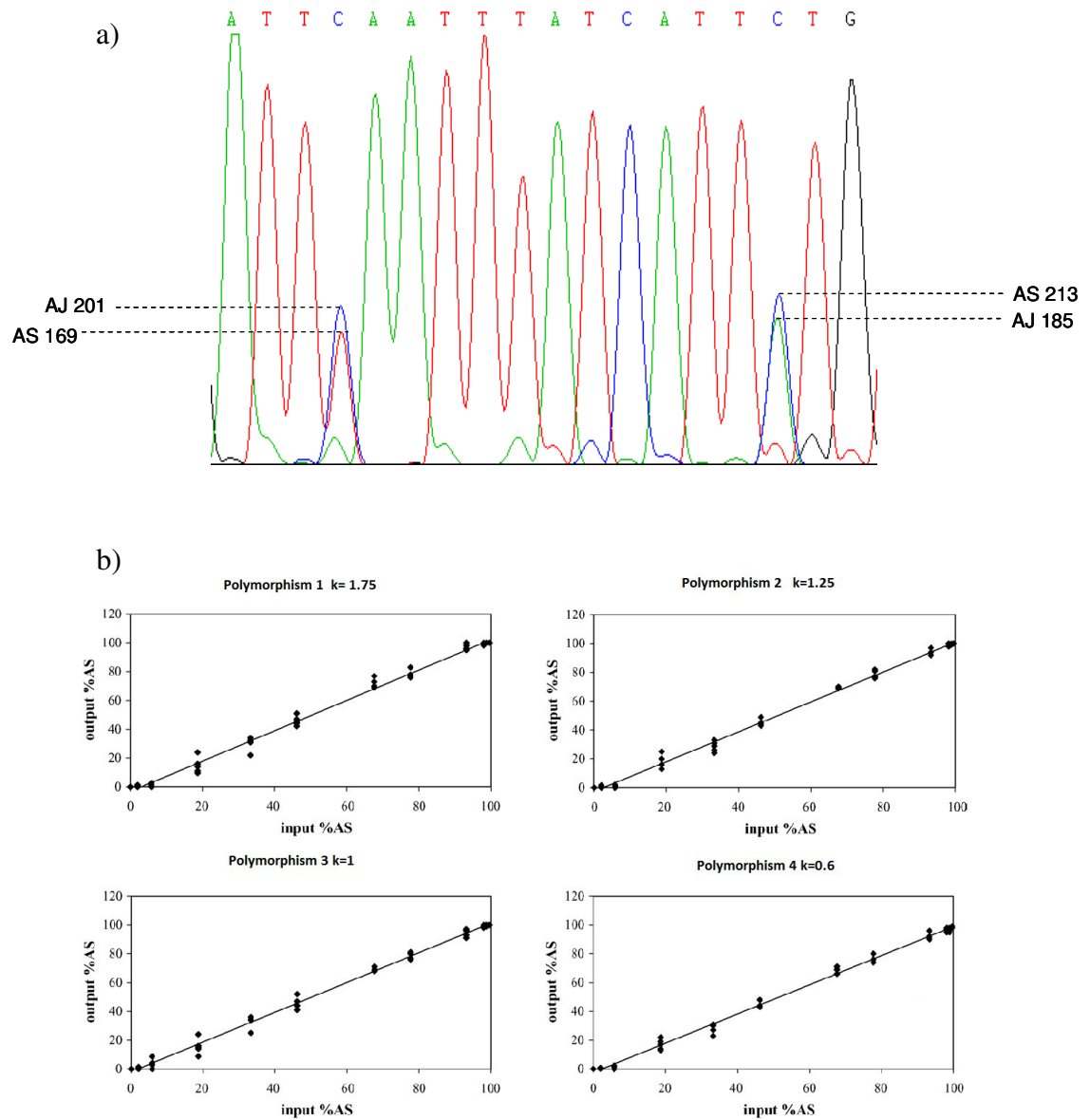


Figure 2.3 The proportional sequencing.

a) The chromatogram representing the fragment of *pcdhps* gene. The two positions polymorphic between the AS and AJ parasite are visible and the measurements of the height of the peaks can be used to estimate the proportions of both strains in the mixed infection

b) The calibration data obtained for the four polymorphisms in *pcdhps* gene (according to Hunt et al., 2005). The x axis on each of the graphs represents the proportion of AS strain in prepared mixtures, obtained by real-time quantitative PCR. The y axis represents the proportional sequencing measurement using one of *pcdhps* polymorphisms. Each of the four repeat of this measurement was plotted. The k_p coefficient for each polymorphism is given.

2.4. Results

2.4.1. AS-30CQ+AJ competition experiment

2.4.1.1. Single infection controls.

Two mice were infected with either of the two strains (AS-30CQ and AJ) and grown either untreated or treated with 5 mg of CQ kg⁻¹ day⁻¹ for 16 days. The parasitaemias of all infections was monitored daily (Fig. 2.4).

Untreated controls. In an untreated mouse, AS-30CQ parasites were observed first at day 5 p.i. Parasitaemia increased until it reached a peak of ~30 % on day 9 and started to recede, disappearing completely by day 13. No parasites were observed for the remaining time of the experiment.

In a mouse infected with the sensitive AJ strain, parasites grew noticeably faster than those of the AS-30CQ strain. The parasitaemia appeared one day earlier (on day 4) and by day 8 it reached 37.8%. At that point of the infection the mouse started showing symptoms of severe disease and had to be sacrificed.

Treated controls. The treated mouse infected with highly resistant AS-30CQ strain developed parasitaemia. However it was significantly delayed in comparison to untreated infections (parasites did not appear till day 11) and the peak of parasitaemia reached on day 14 was only 8.9%.

No parasites were observed in CQ-treated animal, infected with sensitive AJ strain.

Together, these data confirm that both AS-30CQ and AJ were viable and that their drug responses were as expected; sensitive for AJ, resistant for AS-30CQ.

2.4.1.2. Parasitaemia in mixed infections

Five groups of 3 mice were infected with 50:50 mixtures of AS-30CQ and AJ and treated with CQ at various doses for three, four or five days as shown in Fig. 2.1. All animals developed infections. The parasitaemias of all animals were monitored daily during the course of experiments and are shown in Figures 2.5 and 2.6 (violet bars).

In *untreated* mice, parasitaemias developed similarly to the single infection with AJ strain, appearing on day 4 and quickly reaching an average parasitaemia of $43.1 \pm 1.6\%$ on day 8 at which point mice were disposed of (Fig 2.5).

In animals treated with $5 \text{ mg of CQ kg}^{-1} \text{ day}^{-1}$ (Fig 2.6 a-c) the infection was similar to the AS-30CQ treated with the same concentration – the parasites appeared later (day 9 or 10) and reached peak of parasitaemia was lower ($9.9 \pm 2.3\%$) and delayed (day 12/13) in comparison with the previous group..

The mice treated with $10, 20, 30 \text{ mg of CQ kg}^{-1} \text{ day}^{-1}$ also developed infections (Fig 2.6 d-l). They were further delayed in time in comparison with $5 \text{ mg of CQ kg}^{-1} \text{ day}^{-1}$ group (peaks of parasitaemia on average on day 14, 15 and 16 respectively). However, no significant decrease of parasitaemia was observed– the maximum reached was $8.9 \pm 1.5\%$, $10.7 \pm 4.4\%$ and $9.6 \pm 1.7\%$ for 10, 20 and 30 mg of CQ $\text{kg}^{-1} \text{ day}^{-1}$.

All groups contained mice treated for 3, 4 or 5 days. As expected, the number of days of treatment seemed to decrease the peak parasitaemia (the peaks parasitaemias in animals treated for 3, 4 and 5 days were $12.6 \pm 0.38\%$, 9.0 ± 3.0 and 7.8 ± 2.4 respectively) and delay the infection (on average the parasites in animals treated for 3 days appeared a day or two earlier than in those treated for 5 days).

2.4.1.3. Proportions of parasites in mixed infections

All mixed infections consisted of an inoculum containing an equal proportion of both strains AS and AJ. Proportional sequencing was used to confirm this proportion and to monitor the changes of proportion throughout the infection. The results are presented on Fig. 2.5 and Fig. 2.6 (green lines).

The proportion of the AS parasites in the original inoculum was $47.2 \pm 0.6\%$ which suggests that the inoculum was accurately prepared. Later, however, this proportion underwent significant changes in all the groups.

In *untreated animals* the proportion of resistant (AS-30CQ) parasites dropped dramatically. From the initial $47.2 \pm 0.6\%$ it was reduced to $21.3 \pm 0.9\%$ on day 4 and 3.9 ± 1.2 on day 6. By the day 8 the resistant parasites could no longer be detected.

In *treated animals* initially the more resistant AS-30CQ dominated the early part of the infections, as expected. Almost all samples taken on day 7 (no reading could be obtained from the earlier samples possibly due to the very low number of parasites) showed 100% AS in the infections and this proportion was maintained at least until the peak of parasitaemia. The only exception was an animal treated with 20 mg of CQ kg⁻¹ day⁻¹ for 5 days (Fig 2.6-i) which showed 41.2 % of resistant strain on day 6 and subsequent gradual increase of proportion of AS-30CQ till it reached the 100% on day 16 (coincidentally the parasitaemia of this animal was the most delayed and lowest of all infections).

Later in the infection however, some animals showed a decrease of proportion of AS-30CQ. The sensitive strain (AJ) had re-emerged and again constituted a significant part of infection. This re-emergence seems to be correlated with low doses and short time of treatment – it was observed in all animals treated with 5 mg of CQ kg⁻¹ day⁻¹ (the lowest used dose of CQ) and in animals treated with 10 and 30 mg of CQ kg⁻¹ day⁻¹ for 3 days (the shortest time of treatment). It always appeared after the peak of parasitaemia.

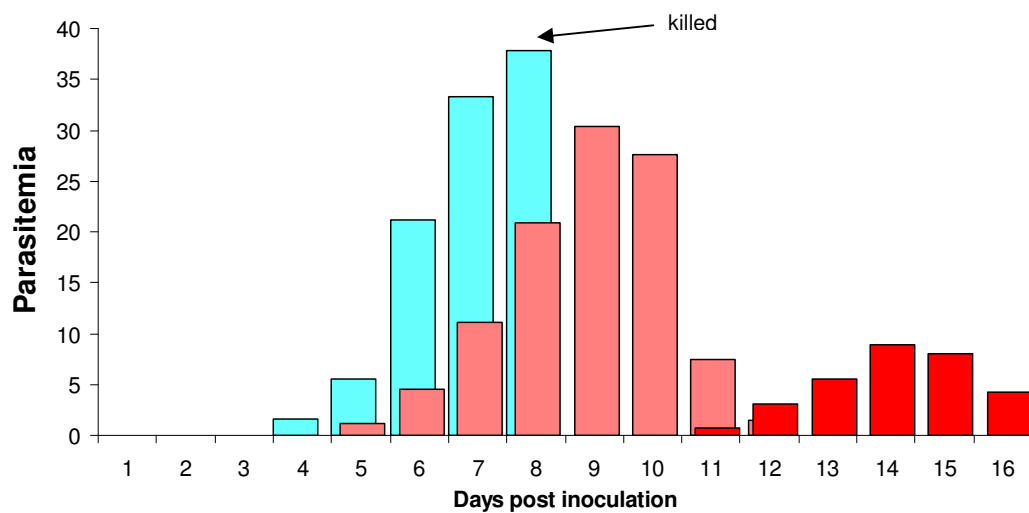


Figure 2.4 The controls for AS-30CQ + AJ infections.

The bars are representing the parasitaemia of ■ - the untreated AJ infection, ■ - the AJ infection treated with 5 mg of CQ kg⁻¹ day⁻¹, ■ - the untreated AS-30CQ infection, ■ - the AS-30CQ infection treated with 5 mg of CQ kg⁻¹ day⁻¹

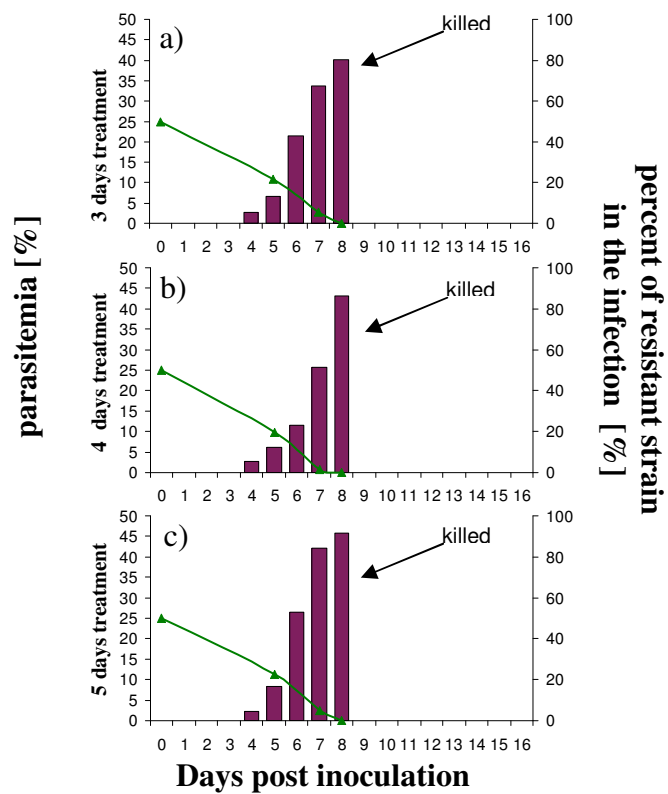


Figure 2.5 The mixed infections with AS-30CQ and AJ strains in untreated animals.

The changes in parasitaemia (violet bars) and proportion of resistant strain in the mixture (green line) during the course infection are recorded. Each graph is representing one animal.

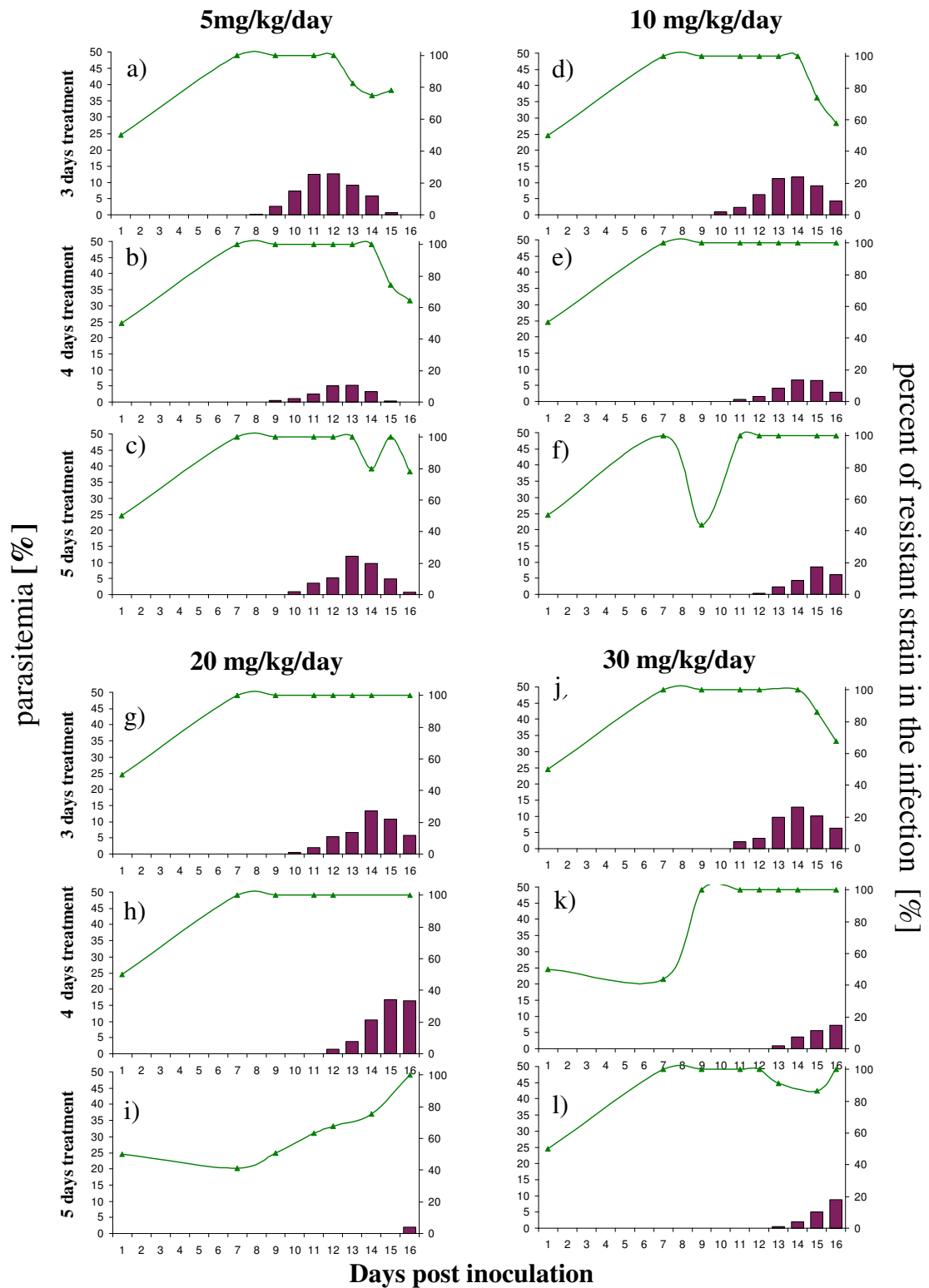


Figure 2.6 The mixed infections with AS-30CQ and AJ strains

Treated with 5 (a-c), 10 (d-f), 20 (g-i) and 30 (j-k) mg of CQ $\text{kg}^{-1} \text{day}^{-1}$. In each group mice have been treated 3 (a,d,g,j), 4 (b,e,h,k) or 5 (c,f,i,l) days. The changes in parasitaemia (violet bars) and proportion of resistant strain in the mixture (green line) during the course of infection are recorded. Each graph is representing one animal

2.4.2. The AS-3CQ +AJ competition experiment

2.4.2.1. Single infection controls

Analogously to the previous experiment, two control infections were established for each of the strains; the low resistant AS-3CQ and the sensitive AJ. One of the mice was left untreated and another was treated with 3 mg of CQ kg⁻¹ day⁻¹ for 3 days. The parasitaemia of all infections is presented on (Fig. 2.7).

Untreated controls. The mouse infected with AS-3CQ strain developed parasitaemia on day 4 and reached the peak of infection on day 9 with 41.5% of parasitaemia followed by sharp decrease. Compared to the highly resistant AS-30CQ strain from previous experiment, AS-3CQ seems to grow quicker as the parasitaemia appears earlier and its peak is higher.

In the animal infected with the sensitive AJ strain parasites appeared on day 4. By day 7 the peak parasitaemia was 43.1% and, because of the symptoms of disease the animal was sacrificed.

Treated controls. An animal infected with AS-3CQ and treated with 3 mg of CQ kg⁻¹ day⁻¹ developed parasitaemia on day 12 and reached its peak day 18 with 20.4%. It presents a significant delay in comparison with AS-30CQ from previous experiment, which managed to develop infection 4 days earlier when treated with higher concentration of CQ (5 mg of CQ kg⁻¹ day⁻¹). However the peak of parasitaemia AS-30CQ reached was only half of the one of AS-3CQ.

The sensitive AJ strain failed to develop measurable parasitaemia during the time of experiment.

Together, these data confirm that both AS-3CQ and AJ were viable and that their drug responses were as expected; sensitive for AJ, resistant for AS-3CQ. It also confirms that AJ grows faster than both AS clones. However no valid comparison can be made between the two AS strains as they were characterised in different experiments, using slightly different treatment regimens and with only one animal per treatment.

2.4.2.2.Parasitaemia in mixed infections

Five groups of mice were infected with a mixture of AS-3CQ and AJ and treated as shown on Fig 2.2 The parasitaemia of all infections was monitored daily is presented as green bars on Fig. 2.8-2.11.

All *untreated animals* (Fig 2.8) developed infections with an average peak parasitaemia of $42 \pm 3.8\%$ on day 7. Afterwards the mice started to show the symptoms of illness and had to be killed on day 10.

The animals treated with $1.5 \text{ mg of CQ kg}^{-1} \text{ day}^{-1}$ - the lowest concentration - also developed infections (Fig 2.9). However, even this very low treatment significantly delayed the parasitaemia – the first parasites were not observed until day 13 and the peak of the infection was on day 17-18. The obtained parasitaemia was also lower – the maximum was $22.3 \pm 3.1\%$ which was around half of the values obtained in untreated infection.

In the group treated with $3 \text{ mg of CQ kg}^{-1} \text{ day}^{-1}$ (Fig 2.10) only two animals out of four developed a detectable parasitaemia during the time of experiment (Fig 2.10 b and c). In both cases the maximal values of parasitaemia (35.6% and 15.3% respectively) were observed on the last day of experiment (day 19).

For two groups treated with the highest concentrations ($5 \text{ and } 20 \text{ mg of CQ kg}^{-1} \text{ day}^{-1}$) no parasites could be detected on the blood smears during the course of experiment suggesting that the drug inhibits both sensitive and resistant strain. However in the $5 \text{ mg of CQ kg}^{-1} \text{ day}^{-1}$ (Figure 2.11) a PCR product could be obtained which can be explained by the presence of very low numbers of parasites in mouse blood, undetectable by microscopy.

2.4.2.3.Proportions of parasites in mixed infection

As in the previous experiment, all animals were inoculated with a 50:50 mixture of both AJ and AS strains and the proportion of parasites was followed using the proportional sequencing of *pcdhps* SNPs. The obtained data are presented as orange lines on Fig. 2.8-2.11.

In *untreated infections* (Fig 2.8) there was an initial decrease in the proportion of the resistant AS-3CQ strain. On day 3 the average proportion of AS-3CQ was

29.6 ± 1% which means that majority of infection was composed from the sensitive AJ strain. However later instead of further decreasing (as observed in AS-30CQ experiment), the proportion of AS-3CQ stabilised itself at a level of ~20% for the remainder of the infection – the lowest observed value was 16.2% and the highest 26%.

In the **1.5 mg of CQ kg⁻¹ day⁻¹** group (Fig 2.9) initially the resistant strain dominated the infection and achieved the proportion of 100% on day 4, when first blood samples were taken. In two of the animals (Fig 2.9 a, d) this proportion was maintained until the end of the experiment. In the other two, however, a gradual decrease in the proportion of the resistant strain was observed. In mouse no 2 (Fig 2.9 b)) this decline started on day 14 and on day 18 its proportion in the infection was 50.35% - similar to the one in the initial inoculum. In animal n° 3 (Fig 2.9 c)) the process started on day 9 and by day 19 the proportion of resistant AS-3CQ strain was only 13.4%. In both cases by the time the peak of parasitaemia was reached, the sensitive parasites constituted the major part of the infection

In the case of **3 mg of CQ kg⁻¹ day⁻¹** (Fig 2.10) the resistant strain AS-3CQ was selected quickly and reached 100% in all infections. However, similarly to the previous concentration, a drop of AS proportion in one mouse suggested the re-emergence of the sensitive AJ strain (Fig 2.10 a) is possible. Interestingly this animal was the one that developed the highest parasitaemia and the only one to reach a peak of infection during the course of experiment.

In the **5 mg of CQ kg⁻¹ day⁻¹** group (Fig 2.11), results of proportional sequencing were obtained despite the lack of detectable parasitaemia during the course of an experiment. The interpretation of this data was difficult due to lower quality of sequences (most likely caused by the small amount of genetic material). However, it suggests the complete selection for resistant strain with the possibility of re-emergence of the sensitive strain later in the infection.

In a last group treated with **20 mg of CQ kg⁻¹ day⁻¹** (the other group in which no parasitaemia was observed) no proportional sequencing results could be obtained either due to lack of PCR product or very low quality, unreadable sequences. It

suggests that the use of this drug concentration reduced the numbers of parasites in blood below the level of PCR detection.

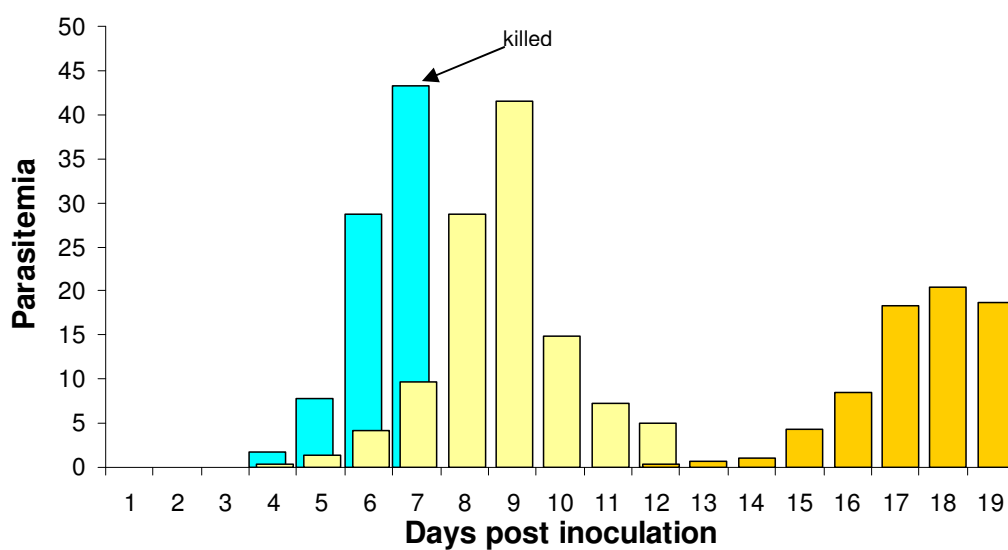


Figure 2.7 The controls for AS-3CQ + AJ infections

The bars are representing the parasitaemia of ■ - the untreated AJ infection, ■ - the AJ infection treated with 3 mg of CQ kg⁻¹ day⁻¹, ■ - the untreated AS-3CQ infection, ■ - the AS-3CQ infection treated with 3 mg of CQ kg⁻¹ day⁻¹

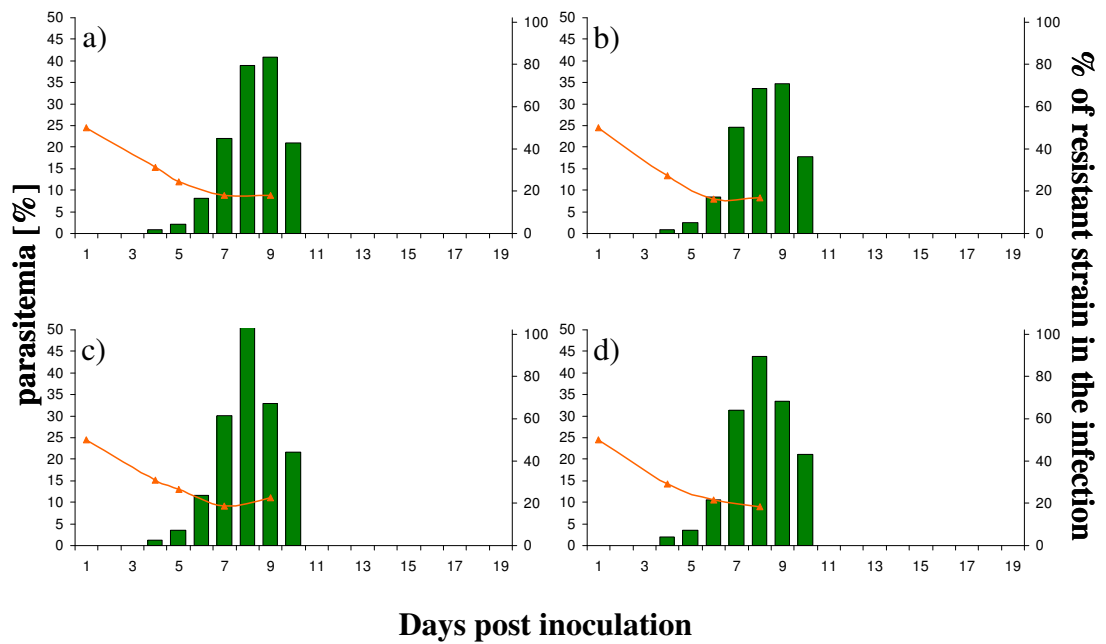


Figure 2.8 The mixed infections with AS-3-CQ and AJ strains treated with 0 mg of CQ kg⁻¹ day⁻¹. The changes in parasitaemia (green bars) and proportion of resistant strain in the mixture (orange line) during the course infection are recorded. Each graph represents one animal.

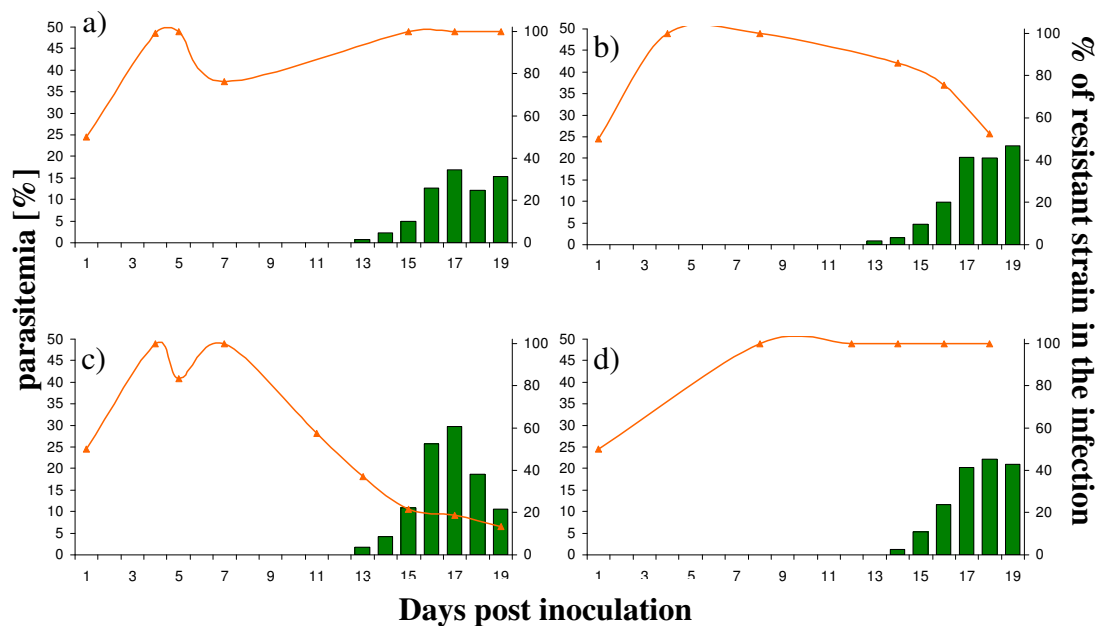


Figure 2.9 The mixed infections with AS-3-CQ and AJ strains treated with 1.5 mg of CQ kg⁻¹ day⁻¹. The changes in parasitaemia (green bars) and proportion of resistant strain in the mixture (orange line) during the course infection are recorded. Each graph represents one animal.

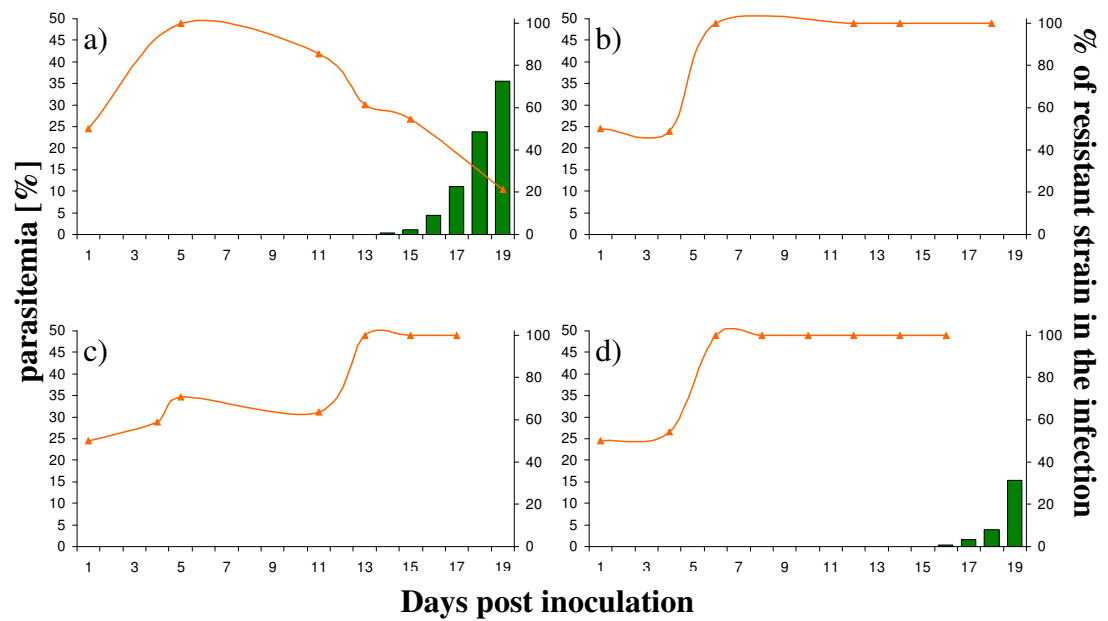


Figure 2.10 The mixed infections with AS-3-CQ and AJ strains treated with 3 mg of CQ kg⁻¹ day⁻¹. The changes in parasitaemia (green bars) and proportion of resistant strain in the mixture (orange line) during the course infection are recorded. Each graph represents one animal.

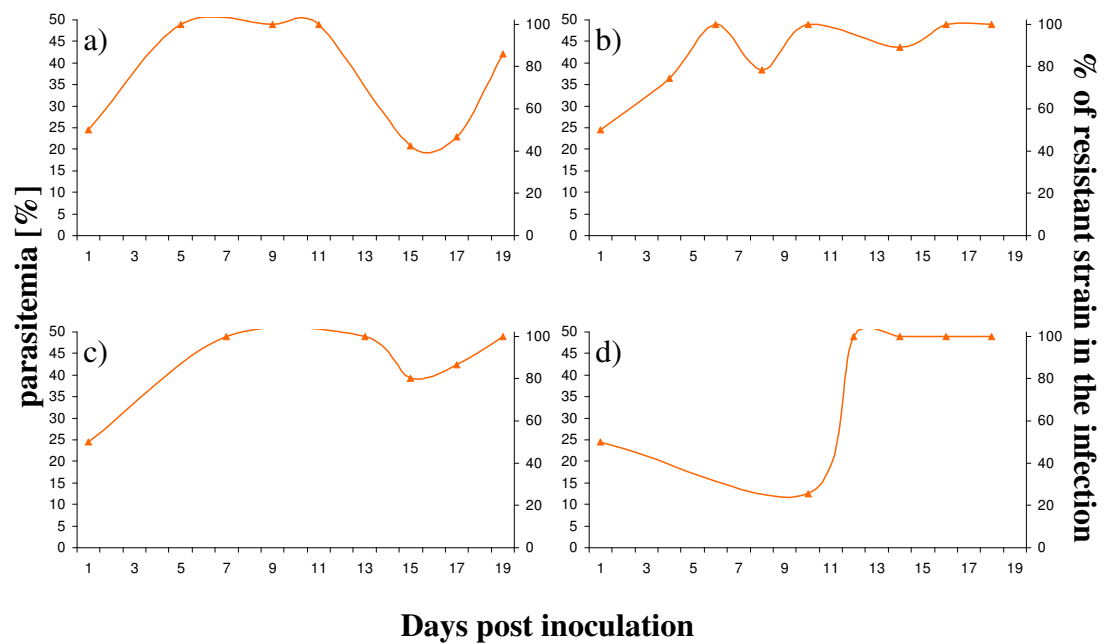


Figure 2.11 The mixed infections with AS-3-CQ and AJ strains treated with 5 mg of CQ kg⁻¹ day⁻¹. The changes in parasitaemia (green bars) and proportion of resistant strain in the mixture (orange line) during the course infection are recorded. Each graph represents one animal

2.5. Discussion

2.5.1. The confirmation of the phenotypes of studied strains

The behaviours of the studied strains were consistent with previous studies (Rosario, 1976, Padua, 1981) confirming the sensitive, low resistant and highly resistant status of the three strains (AJ, AS-3CQ and AS-30CQ, respectively) used in the present study.

The sensitive AJ strain presented a good pattern of growth in the untreated controls and was present in significant proportions in all the untreated mixed infections. No growth was observed in any of the treated single infection controls and in mixed infection the chloroquine treatment invariably resulted in infections composed in 100% from the resistant strain.

Interestingly the cases of recrudescence suggest that the AJ parasite is present in some of the treated infections in low numbers. It is possible that in favourable conditions, such as the reduced CQ plasma concentration following CQ administration, AJ can resume growth. Further research is required whether this phenomenon is characteristic for mixed infections only (no recrudescence was observed in any of the controls).

The low resistant AS-3CQ produces good infections in untreated animals. It can also produce moderate but delayed parasitaemias at lower drug concentrations (1.5 and 3 mg of CQ kg⁻¹ day⁻¹). This parasite is, however, completely inhibited by higher concentrations (5 and 20 mg of CQ kg⁻¹ day⁻¹).

Finally the highly resistant AS-30CQ was able to grow after all the CQ treatments used here (even under 30 mg of CQ kg⁻¹ day⁻¹ for 5 days) although the higher concentrations reduced and delayed its growth, relative to untreated controls.

2.5.2. The drugging regime to select the highly resistant recombinants from low resistant and sensitive ones

The AS-30CQ+AJ experiment showed that all treatments (5, 10, 20 and 30 mg of CQ kg⁻¹ day⁻¹) offer complete selection (varying growth of AS-30CQ, no growth of AJ) ensuring that the infection consisted exclusively of the AS-30CQ resistant strain until the peak parasitaemia had passed. Moreover, the AS-30CQ+ AJ experiment indicated that 5 mg of CQ kg⁻¹ day⁻¹ (and higher doses) are enough to inhibit also the low resistant clone, AS-3CQ. However, all concentrations used also significantly delay and decrease the parasitaemia potentially generating less material for genetic analysis. The lower (5, 10 mg of CQ kg⁻¹ day⁻¹) treatments are, therefore, preferred as these negative effects are lower, although the higher ones may also be useful as they are likely to provide a stronger selection.

As the number of days of drugging doesn't seem to affect the strength of the selection but only the time of re-emergence of sensitive parasites and appearance of peak of parasitaemia, the shortest drugging regime has been chosen and used in all following experiments.

As some cases of AJ recrudescence were observed after the peak of parasitaemia in order to ensure the optimal selection the parasites were need be harvested before it occurs. As the peak of parasitaemia can be determined only retrospectively and can vary great, for practical purposes it was decided that the parasites are going to be harvested once the infected animals reached 5% of parasitaemia which always seem to occur a day or two before the peak.

2.5.3. The drugging regime to select all resistant recombinants from sensitive ones

Initially all of the concentrations used selected for the resistant strain. However, only in the lowest of them (1.5 and in one case 3 mg of CQ kg⁻¹ day⁻¹) did the parasites reach a detectable peak parasitaemia during the course of experiment. Others reduced parasitaemia below the detectable threshold and, therefore, were discarded as potential treatment regimes. Moreover, in the majority of the mice that

developed acceptable parasitaemia, the re-emergence of the sensitive AJ strain was observed very early and by the time parasites could be observed, AJ was reaching a significant percentage of infection. It makes selection of low resistant recombinants of the cross difficult as they are likely to be outgrown either by highly resistant recombinants (at the higher drug concentrations) or by the sensitive one (at lower drug concentrations). 1.5 mg of CQ kg⁻¹ day⁻¹ is the only concentration that offers relatively good (although not complete) selection for resistant parasites when the parasitaemia can be observed. Therefore I decided to use it for the selections, taking into account that the results may be distorted by re-emerging sensitive strain.

2.5.4. The fast-growing sensitive parent: implications for the cross

The sensitive parent AJ used in subsequent genetic crosses has, in untreated animals, consistently exhibited higher growth rates than the resistant AS parasites (AS-3CQ and AS-30CQ). It may be caused by the general differences between AS and AJ genotypes. The previous experiments (de Roode et al., 2005) comparing the growth of various strains selected from AS-sens and AS-pyr parasite (the progenitors of AS lineage, Fig. 1.7.2) to AJ indicated that, despite the observed variation between various AS clones, they were overall characterised by the lower virulence than AJ strain (both in single and mixed infection). However, it is also possible that the observed differences are caused by the negative effects of mutations conferring the CQ resistance in the AS lineage (as discussed in the section 1.3.3.)

This feature of sensitive parent may have a strong impact on prospective crosses. The higher growth rate of this sensitive parental strain means that the highly resistant recombinants would be at great disadvantage when the cross progeny is grown without drug selection if they have inherited the AS alleles. It could significantly decrease the numbers of resistant recombinants (or even eliminate them completely if the growth disadvantage is due directly to resistant mutations) during the drug free passages that take place before the selections. It has to be taken into account during the analysis and interpretation of LGS results (see section 3.5 of Chapter 3).

2.5.5. Possible fitness cost of chloroquine resistance in *P.chabaudi*

It was already mentioned that the fitness cost of CQ resistance may be one of the factors contributing to the slower growth rate of the resistant parasites compared to the AJ strain. However further experiments (like the competition between the CQ sensitive and CQ resistant strain from the lineage) would be required to confirm this possibility.

Additionally some differences could be observed between the behaviour of the two resistant AS clones in untreated infections. In the ‘single infection controls’ the highly resistant AS-30CQ clone appears one day later and is reaching a peak parasitaemia which is 10% lower than that of the low resistant AS-3CQ. Moreover, during the untreated mixed infection the AS- 30CQ is completely outgrown by sensitive AJ strain within the first seven days (*i.e.* reduced below the threshold of detection of proportional sequencing assay for the remainder of the infection). The low resistant AS-3CQ clone instead, in the presence of the same competitor, is reduced but is still present to a level of at least 20%. These data indicate that the more resistant clone (AS-30CQ) may have a lower fitness than AS-3CQ. As the two clones are congenic it is possible that this difference is caused by the fitness cost of high CQ resistance present in AS-30CQ.

In Chapter 6 I describe further investigations of this phenomenon and a competition experiment between the two resistant strains.

3. Generation and LGS analysis of AS-30CQ x AJ cross

3.1. Introduction

In the next two chapters I describe the use of Linkage Group Selection (LGS) to map the loci related to drug resistance in *Plasmodium chabaudi*. The general description of this method as well as an overview of the experimental design is presented in Chapter 1 section 1.8.1. In summary LGS includes the generation of the cross between the resistant and sensitive parent, the phenotypic selection of the pooled progeny, and the quantitative analysis of the frequencies of genetic markers in derived populations. Therefore before the start of experiment the resistant and sensitive strain, the drug selection regime and quantifiable genetic markers had to be decided upon.

3.1.1. The resistant and sensitive strain

Two strains with different levels of drug resistance were required for the genetic cross.

AS-30CQ, a clone belonging to the *P.chabaudi* AS lineage (described in Chapter 1 section 1.7.2), was chosen as the drug-resistant parent. Being the final product of three rounds of CQ selection (Padua, 1981); it must contain all of the mutations that contribute to the highest level of resistance in the lineage. Therefore it can be used to map all of the loci involved in high-level chloroquine resistance in a single series of experiments. Additionally, recently it has been demonstrated that AS-30CQ is also artemisinin resistant even though no selection for resistance to this drug was applied. Consequently the same genetic cross can be used to investigate the genetic basis of ART resistance as well.

The **AJ**, another *P.chabaudi* strain, genetically different from AS lineage was chosen as the sensitive parent. Its drug response phenotypes are well-characterised: it is sensitive to both CQ and ART. It has previously been used for numerous genetic crosses showing the ability to recombine with the AS genotypes (e.g. Padua, 1981, Carlton 1998, Cravo et al., 2003, Culleton et al., 2005, Hunt *et al.*, 2007). Moreover,

a large number of quantitative genetic markers between AJ and parasites of the AS lineage were available to map the proportion of parental alleles in the progeny (Martinelli et al., 2004 Cheesman et al., 2007).

3.1.2. LGS drug selection conditions

The aim of the LGS experiment described in this chapter was to map all of the loci involved in high-level CQ resistance. Therefore a drug dose selecting the highly resistant genotypes (containing all the loci under selection) from the rest of the population had to be used. The preliminary experiments described in chapter 2 were performed in order to establish the optimal selection conditions. The 3 day (d 0-2 pi) regimes of 5, 10 or 20 mg CQ kg⁻¹ day⁻¹ should select the highly resistant parasites from the low resistant and sensitive ones and were used for cross selection.

3.1.3. Pyrosequencing

A genome-wide genetic analysis of progeny requires the use of quantitative genetic markers, spaced across the genome. Here, a set of pyrosequencing assays was used for that purpose.

Pyrosequencing markers are based on “sequencing by synthesis” approach and can be used to estimate the proportion of SNPs in a mixture, (Fig. 3.1). As each assay is based on the previously identified SNP, in contrast to AFLP or RFLP markers, it can be designed in the chosen fragment of the genome and its location is known without the subsequent mapping. Currently >100 of such assays were designed between the AS lineage and AJ strain and they have been successfully used in the investigations of genetic basis of immunity in *P.chabaudi* (Cheesman et al, 2010).

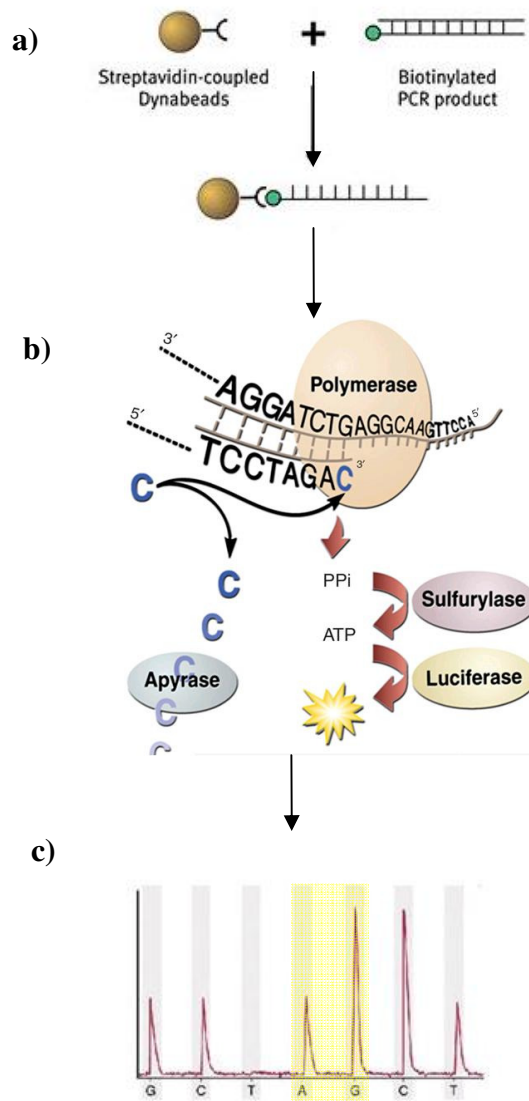


Figure 3.1 The use of pyrosequencing to calculate the proportion of SNP based genetic markers in the population.

A single stranded template containing the polymorphism is prepared by PCR amplification with biotinylated primer, and hybridisation of the product to the streptavidin beads (a). It is then hybridised with sequencing primer in the presence of an enzymatic complex (including DNA polymerase, ATP sulfurylase, luciferase, and apyrase) and the substrates (adenosine 5' phosphosulfate (APS) and luciferin). To this mix the dideoxy nucleotides are added sequentially (b). The incorporation of each nucleotide by polymerase results in generation of equimolar amounts of free pyrophosphate (PPi) which is a substrate of sulfurylase, that catalyses the generation of adenosine triphosphate (ATP) from APS. This in turn is used in reaction catalysed by luciferase converting luciferin into oxyluciferin and resulting in the emission of light. The cycle is then completed by apyrase metabolising the remaining dideoxynucleotide and ATP. The emitted light is detected by charge coupled device chip and is presented on the output as a peak with the height proportional to the amount of the incorporate nucleotide (c). In each assay the sequence of ~10 nucleotides is added, including the two bases of studied polymorphism. The peaks from the two bases of polymorphisms are used in order to calculate the respective proportion of both bases in the mixture. Other peak provide the control and calibration of the reading.

3.2. The experimental design

3.2.1. The generation of the cross

The design of this experiment is presented schematically in Fig. 3.2. In order to generate the initial cross, a group of 5 mice was infected with 10^5 of a 1:1 mixture of the resistant (AS-30CQ) and the sensitive (AJ) strain. Assuming the similar rates of gametocyte generation in by both strains in mixed infection, random mating and equally efficient mosquitoes transmission rates of both genotypes; this proportions should provide the maximal numbers of recombinant parasites in the cross progeny (50%).

On day 5 post infection, when the mice had developed gametocytes, each animal was fed to a mosquito cage containing ~ 200 *Anopheles stephensi* mosquitoes. Fertilisation, recombination and meiosis took place within the mosquitoes. Booster blood feeds were performed (after 5 days) using uninfected mice, to maximise maturation of parasites.

After 7-8 day post infection, a sample of mosquitoes was dissected to check for the presence of oocysts – the transitory parasite stage - and therefore prove successful transmission. The number of oocysts counted was used to estimate the total number of progeny present in all the mosquitoes. Ten days later, when the oocysts had burst and the sporozoites had migrated to the salivary glands, the glands were collected and used as inoculum to infect another group of 5 mice (such an "artificial" infection is generally more efficient then the natural one (*i.e.* occurring during the mosquitoes feeds). On day 6 p.i. the blood samples were taken from all infected animals and the proportional sequencing of *pcdhps* gene, on chromosome 14, was used to confirm the presence of alleles from both parents in the infection. Then the progeny were recovered and used for further passages and selection. At this step the parasites were frozen down to be used for subsequent selections and backcrosses.

3.2.2. The amplification of the cross progeny

The extra amplification step had to be introduced in order to generate enough progeny to infect all animals for selection. In order to amplify the parasites while

maintaining the diversity of the population, the cross was passaged in two groups of two mice. One group was treated with 5 mg CQ kg⁻¹ day⁻¹ 3 days (d 0-2) and one left untreated. The treatment was used in order to preserve the highly resistant parasites that otherwise might have been eliminated from the cross due to their lower fitness.

When the mice in both groups reached the detectable parasitaemia the parasites were mixed in a 1:1 proportion and used to infect the animals for subsequent selection.

3.2.3. Selection

During the selection, four groups of 4 mice each were infected with 10⁶ of AS-30CQxAJ parasites and treated with 0, 5, 10 or 20 mg CQ kg⁻¹ day⁻¹ respectively. The growth of parasites was monitored and when all parasites from the groups reached > 5% parasitaemia the mice were killed and bled out. The blood from all animals from the same group was pooled and DNA isolated for further genetic analysis.

3.2.4. Genetic analysis

The allele frequencies of the genetic markers in the population were estimated using the AS/AJ pyrosequencing assays and the generated data were analysed to detect selection valleys.

Out of ~100 developed pyrosequencing assays 66 were used for the analysis presented in this chapter (see Appendix C for the detailed list). Initially 3 or 4 markers on each chromosome were chosen in order to assess the overall shape of the genome scan. Later additional assays were added starting with the regions of particular interests (like chr 11, 2 or 7 that have previously been connected with CQ, ART and PYR resistance respectively) with the intention to test all the available markers. The addition of new markers however was stopped after the analysis of initial data indicated that it is unlikely to result in substantial data improvement (see section 3.5.4).

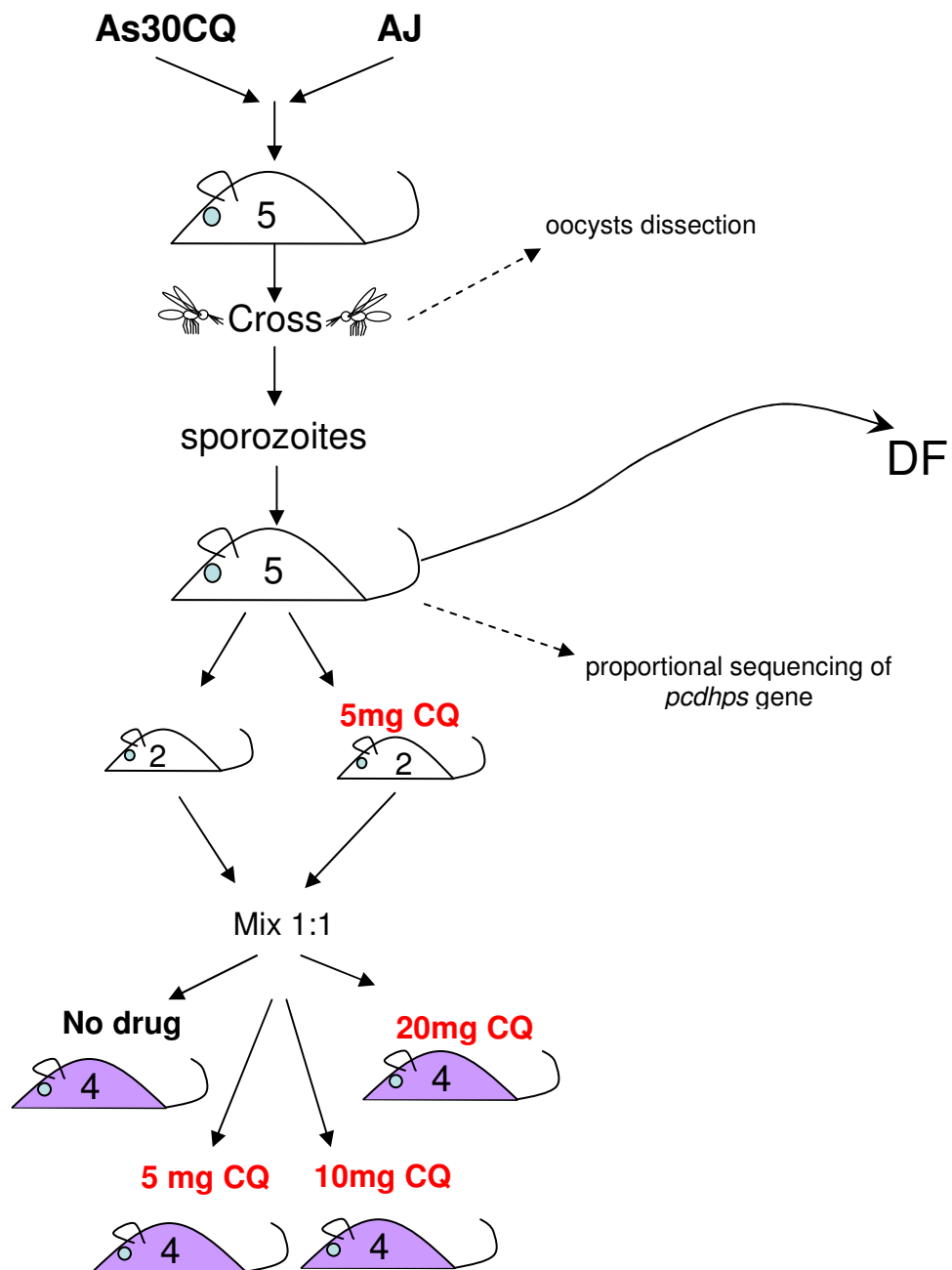


Figure 3.2 The schematic diagram of AS-30CQxAJ LGS experiment

The full arrows represent the transfer of the parasites, the dashed ones – a sample taken for control step. The number inside each mouse represents the number of animals used. The drugging regime used is shown above each mouse. DF- parasites transferred into the deep freeze.

3.3. Materials and methods

3.3.1. Parasites, mice and mosquitoes:

The origins and phenotypes of the resistant AS-30CQ strain and the sensitive AJ have been described elsewhere (Chapter 2, section 2.3.1.).

10-12 week old CBA mice were used for passage and selection of parasites. 10-14 week old C57BL/6 black mice were used for mixed infections for mosquito feeds as this strain is known to develop higher gametocytemias and milder disease symptoms compared to CBA mice (Richard Carter, Paul Hunt, personal communication). All animals were housed and fed as described in section 2.3.1

The mosquitoes for the experiments were taken from the *Anopheles stephensi* colony maintained at the University of Edinburgh under standard insectary conditions; 27°C temperature, 70% humidity, 12h/12h light/dark cycle, and unlimited access to water supplemented with 10% glucose. Routinely, female mosquitoes were allowed to feed on rats and to deposit eggs in eggbowls provided for 2 days, 2 days after each rat feed. Experimental pupae were collected from low-density larval trays and allowed to hatch in experimental cages (~250 pupae per cage). One day before feeding to infected mice, the glucose solution was replaced with water.

3.3.2. Mouse infections

All single and mixed infection were set up, drugged and monitored as described in sections 2.3.2,- 2.3.4. of chapter 2.

3.3.3. Mosquito feeds

On day of feeds the mice were anaesthetised with Rompun/Vetalar mixture and placed on the top of mosquito cages, one mouse per cage. The mosquitoes were allowed to feed for 20 minutes. The mice were killed immediately after feeding. The booster feeds were arranged similarly.

3.3.4. Dissection for oocysts

A sample of 10 mosquitoes from each cage was dissected. The mosquitoes were killed with chloroform and placed under binocular dissecting microscopes.

Two cuts were made between abdomen and thorax using syringes with needles attached. Then the abdomen and thorax were pulled apart. The exposed gut was cut out and suspended in a drop of Ringer solution. These guts were placed on microscope slides, covered, and observed under microscope in order to count oocysts.

3.3.5. The recovery of sporozoites from salivary glands

16-17 days after infection mosquitoes were killed with chloroform and dissected. Needles attached to syringes were used to separate the head and salivary glands attached. The glands were removed, suspended in 50% Ringer/Fetal Calf Serum solution (see Appendix A) and sheared in a glass mortar with Teflon pestle. The resulting suspension was diluted to an appropriate volume (allowing the injection of 100ul per mice) and injected intraperitoneally into a new group of mice

3.3.6. The proportional sequencing of *dhps* gene

As described in section 2.3.5 of chapter 2.

3.3.7. Large scale DNA isolation

In order to isolate the parasite DNA, the mouse blood was put through two columns, containing fibrous cellulose powder, and two Plasmodipur filters (Euro-Diagnostica) in order to remove the mouse white blood cells. The RBC obtained were lysed with 0.15% saponin solution and washed with PBS buffer till the supernatant was clear. The resulting parasite pellets were re-suspended in parasite lysis solution (see Appendix A) and incubated at 37°C overnight. From the lysed solutions, DNA was extracted with standard phenol/chloroform extraction protocol and ethanol precipitation.

3.3.8. Estimation of frequencies of genetic markers using the pyrosequencing

3.3.8.1. Design of pyrosequencing assays

The design of pyrosequencing assays set was a collaborative effort to which many people contributed. My part in this project consisted on design of 7 pyrosequencing assays on chromosome 11 (marked in the table in Appendix C)

Each pyrosequencing assay was based on SNPs identified between the AS and AJ. Because the complete *P. chabaudi* genome was not yet available, the *Plasmodium falciparum* genome (<http://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?id=36329>) and the known syntenic relationships between the *P.falciparum* and rodent malaria (Kooij et al., 2005) was used to place assays in the desired positions.. The following approach was used

- a fragment of *P.falciparum* genome syntenic to the locus of interest in *P.chabaudi* was identified using the syntenic map generated by Kooij et al.,
- a protein encoded by the gene situated within that region was chosen and used to search the translated contig sequences of *P.chabaudi* database (http://www.sanger.ac.uk/cgi-bin/blast/submitblast/p_chabaudi)
- the identified contig' sequence was compared back to the *P.falciparum* protein database in order to confirm the unique match
- DNA sequence in syntenic contigs was used to design SNP finding primers
- a ~1000 bp fragment from the genomes of AS and AJ parasites was sequenced using dideoxy sequencing. The genetic distance between the two strains ensured that within most of the 1000 bp fragments one or more SNPs could normally be identified.
- the discovered polymorphism and surrounding sequence was used to design the pyrosequencing primers used in the PSQ Assay design software (Biotage). Each assay was composed of three primers – two amplifying primers (one of them biotinylated) and one sequencing primer.

Each assay was named according to its putative orthologue position in the *P.falciparum* genome. Pcpfxx-yyyy annotation was used where xx is the chromosome and yyyy is the putative distance on that chromosome in kb (e.g. assay 06-1001 was localised on the fragment of the genome syntenic to the one found 1001 kb from the left along chromosome 6 in *P.falciparum*). The effort was taken to space the designed assays at the regular intervals across the genome and to keep the distances of less than 100 kb between the assays. However it was not always possible because either no suitable polymorphism could be found in the particular region, or no working assay could be designed based on it.

3.3.8.2. Pyrosequencing

The PSQ 96A HS pyrosequencing system from Biotage (currently acquired by QIAGEN, <http://www.pyrosequencing.com>) was used to estimate the proportions of SNP alleles in the progeny. DNA was amplified using the primers designed in previous section (see Appendix C for the complete list), the standard PCR protocol (see Appendix B) and the following PCR program:

Temperature	Time	29 x
95°C	10s	
48°C	1min	
68°C	1min	
94°C	1min	
50°C	30s	
68°C	30s	
68°C	2min	

The amplified PCR product was purified using the streptavidin beads, mixed with the sequencing primer and inserted into the PSQ 96A HS machine according to the manufacturer's instructions. The PSQ_HS_96A software was used to program and control the run. All consumables and reagents used for these parts of the experiment were supplied by Biotage/Qiagen.

In order to minimize the technical variation for each marker, all the sample DNAs and the negative control were genotyped during the same run (same 96 well plate) whenever possible. All analysis was performed in triplicate using independently amplified products.

3.3.8.3. Data analysis

The three replicates for each assay were averaged. As a general guideline, when the standard error of the mean was $> \sim 5\%$ of the mean itself, the assay was repeated and if no improvement showed, discarded from the analysis

The comparative index (CI_i) was calculated for each SNP marker by dividing the AJ allele frequency at assay i in the treated population ($treated_i$) by the AJ allele frequency at the same SNP in the untreated population.

$$CI_i = treated_i / untreated_i$$

The comparative index is therefore a measure of the change in allele frequency of a marker between the treated and untreated population.

The marker frequencies and/or CI were plotted against their cumulative genomic position in *P.chabaudi*. Initially it was calculated using the distances between the markers estimated according to *P.falciparum* genome with extra 50 kb added at the end of each chromosome in order to account for *P.chabaudi* specific non syntenic regions. Later after the publication of assembled *P.chabaudi* genome the distances between the markers were adjusted according to *P.chabaudi* data. The resulting genome scans were analysed in order to identify the potential selection valleys.

3.4. Results

3.4.1. Generation of the AS-30CQxAJ cross

The 5 mice carrying mixed infections of resistant and sensitive strains were used to infect mosquitoes in 5 cages. Oocysts were detected in three cages (Table 3.1). The two cages with the highest oocysts count per mosquito (Cage 1 and 2, with means of 16.4 and 3.45, respectively) were chosen and the salivary glands from all the females (73 from first cage and 55 from the second) were dissected and used to infect the new group of 5 mice.

Once the infections developed in all 5 animals, proportions of *pcdhps* AS/AJ alleles in the recovered progeny were estimated, using proportional sequencing. The results are shown in Table 3.2. No PCR product could be obtained from one of the mice. The proportion of AJ alleles in the other four cages lay between 52% - 67%. These data indicate the presence of both parental alleles in these infections.

Mosquito number	Cage 1	Cage2	Cage 3	Cage4	Cage5
1	10	~50	0	0	0
2	1	12	0	0	0
3	3	2	1	0	0
4	4	0	0	0	0
5	0	0	0	0	0
6	0	0	3	0	0
7	9	0	7	0	0
8	11	0	4	0	0
9	0	~55	7	0	0
10	0	~45	2	0	0
Average	3.8	16.4	2.4	0	0
Nr of mosquitoes dissected	73	55			
Potential number of oocysts obtained	277.4	902			

Table 3.1 The results of mosquito dissections for AS-30CQxAJ cross.

The number of oocysts found in each mosquito is given. When the number of oocysts was too large for the exact count, an estimate (~) is given. The results from the cages chosen for salivary glands dissections are shaded.

Mouse	Proportion of AJ <i>pcdhps</i> allele
1	57%
2	66.9 %
3	52.2%
4	No PCR product
5	67.4%

Table 3.2 The results of proportional sequencing of recovered progeny from AS-30CQxAJ cross.

3.4.2. Selection of the cross

In order to ensure that the LGS analysis was carried out on a population containing both CQS and CQR parasites, the propagated parasites were passaged in either the presence or absence of 5 mg CQ kg⁻¹ day⁻¹ and re-mixed. These pooled parasites were then grown in groups of 4 mice treated with 0, 5, 10 and 20 mg CQ kg⁻¹ day⁻¹ in order to select for the high CQ resistant parasites. The parasitaemias in each group were monitored and the results are shown in Fig 3.3.

In the untreated animals the parasites exhibited the fastest growth appearing on day 3 and reaching over 20% parasitaemia on day 6. All the treated mice also developed parasitaemia. However, parasites appeared much later than in the untreated group and this delay was correlated with the concentration of drug used (the first parasites were observed on days 10, 12 and 13 for 5, 10 and 20 mg CQ kg⁻¹ day⁻¹ group respectively). These growth curves were very similar to those observed in AS-30CQ+AJ mixtures during the experiment described in Chapter 2 (Fig. 2.5 and 2.6 Chapter 2).

Parasites were harvested when all mice from each group had reached >5% parasitaemia which was on day 6 (untreated), 13 (5 mg CQ kg⁻¹ day⁻¹ group), 15 (10 mg CQ kg⁻¹ day⁻¹ group) and 16 (20 mg CQ kg⁻¹ day⁻¹ group).

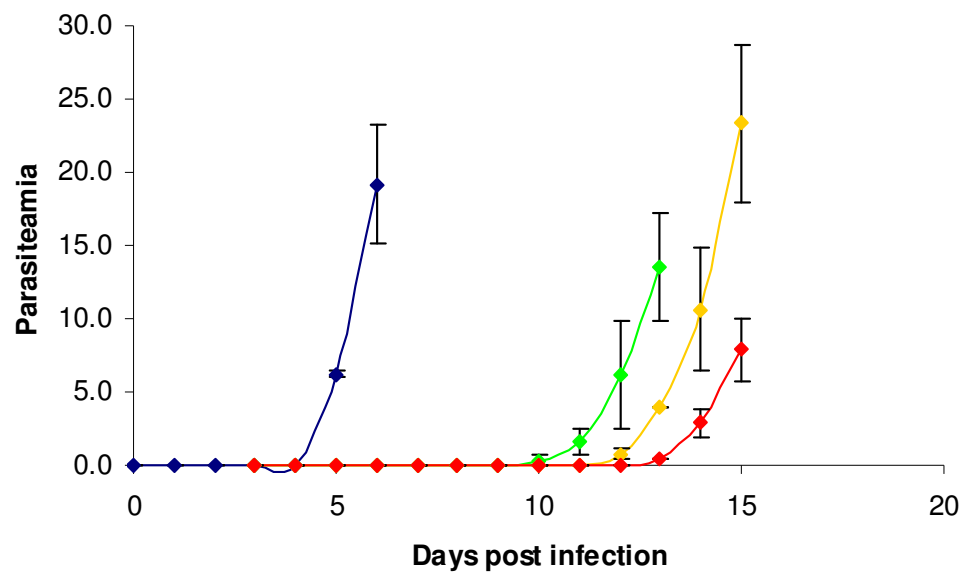


Figure 3.3 The parasitaemia of AS-30CQxAJ cross

The parasitaemia of groups animals untreated (■) and treated with 5 (■), 10 (■) and 20 (■) mg CQ kg⁻¹ day⁻¹.

3.4.3. The quantitative genetic analysis of the initial cross

Having isolated the DNA, the frequencies of parental alleles at 66 SNP markers distributed across the genome were estimated. The resulting genome scans are presented in figure 3.4.

3.4.3.1. Untreated population

In the untreated population we observe very high frequencies of AJ alleles across the whole genome (Fig. 3.4a). The average percentage of the sensitive (AJ) parental alleles was 89.9% (range 77.5 – 100%). Overall, the distribution of parental alleles across the genome is very uniform – no selection valleys were observed. This distribution (~90% AJ) contrasts with estimations of AJ *pcdhps* allele frequency in the untreated progeny arising from the sporozoite infections in which more even proportions of both strains were observed (Table 3.2). Therefore the parasites containing mainly AS alleles appear to have been lost during the growth in untreated animals.

3.4.3.2. Treated populations – general picture

All the treated populations gave similar genome scans to each other but strikingly different to that arising from the untreated population (Fig 3.4 b-d).

Overall, the genome-wide mean percentage of AJ alleles in the treated mixtures was strongly decreased in comparison to the untreated population, being on average 46.13% for 5 mg CQ kg⁻¹ day⁻¹, 44.06% for 10 mg CQ kg⁻¹ day⁻¹ and 45.18% for 20 mg CQ kg⁻¹ day⁻¹. These data suggest that, in general, AS alleles are being selected, or at least maintained, by the drug.

The dominant feature of these scans is the extremely bimodal distribution of parental allele proportions: *i.e.* markers have either a very high or a very low AJ proportion with few markers showing intermediate values. For example, in all three scans, markers on chr 7 all show high AJ proportions (>90 % AJ) while those on chr 8 are universally decreased (<25% AJ). Furthermore, there are many instances where a bimodal distribution can be observed *within* a chromosome. For example, chr 1, 2, 6, 9, 11 and 13 all contain markers with a strong bias towards either AS or

AJ alleles. This effect seems to be the strongest at 10 mg CQ kg⁻¹ day⁻¹ treatment (when no markers with a value between 20 and 80% AJ allele can be observed) and weakest at 5 mg CQ kg⁻¹ day⁻¹.

Overall, these data give the appearance that whilst drug treatment appears to have decreased the total proportion of AJ alleles (genome-wide), individual AJ markers tend to be close to either 0 or 100%. Many regions of the genome therefore appear to show heavy selection (either positive or negative).

3.4.3.3. Treated population – potential loci of CQ selection?

In total around 60% of the tested markers show the reduction of AJ allele between the untreated and treated population – a sign of potential CQ selection. On at least five chromosomes (3, 5, 8, 11 and 12) all markers are consistently decreased with all three treatments and many single markers on the other chromosomes are also strongly reduced. In some cases the group of decreased markers forms a shape that could be interpreted as selection valley (*e.g.* chr 5, 8 and 11). More often however, the small number of markers and/or sudden changes of frequency (*e.g.* chr 13) do not allow a definitive interpretation.

Interestingly in all the cases, the selection seems to be the strongest at 10mg/kg/day of CQ and not at the highest concentration.

The possible interpretations of these results are now discussed.

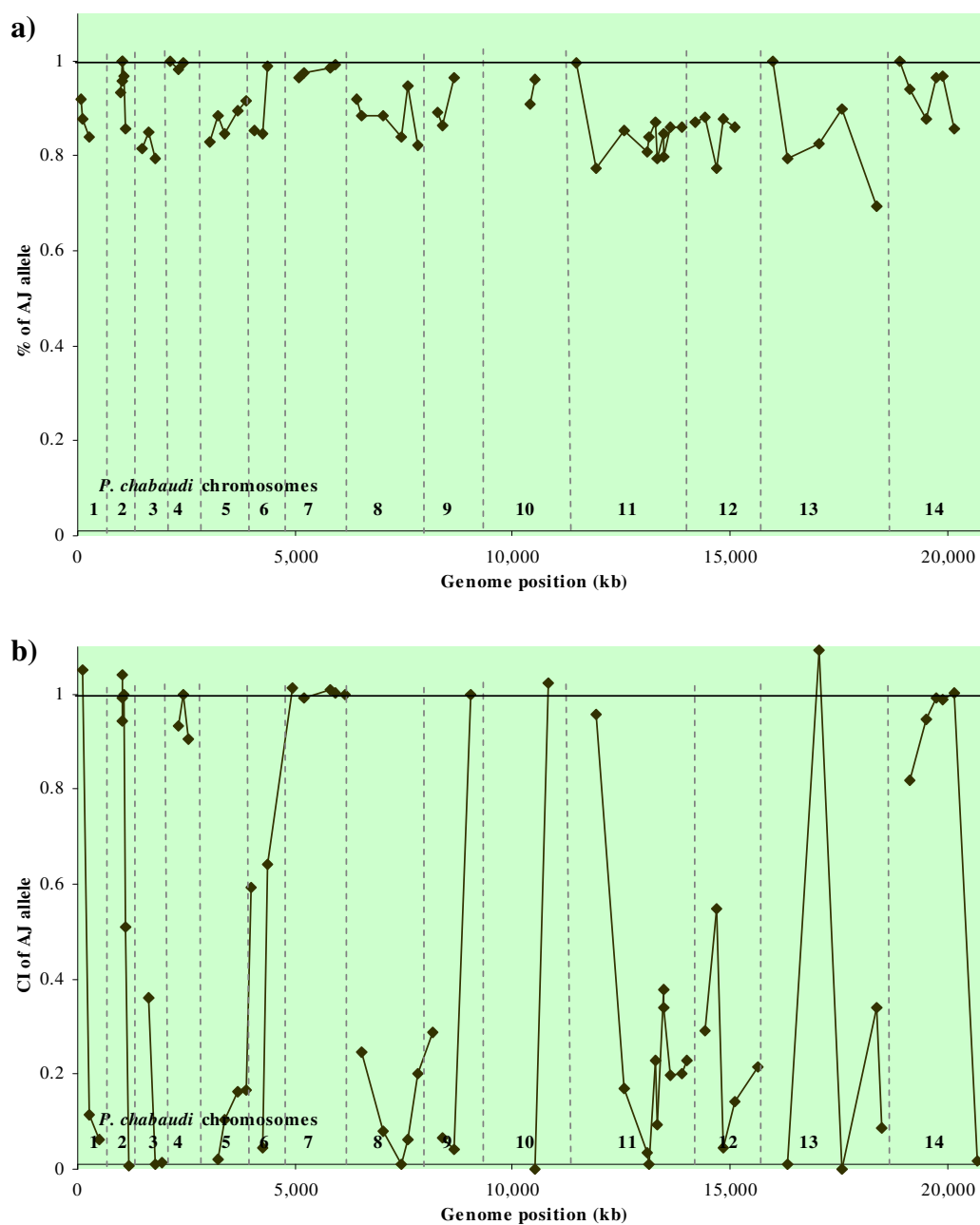


Figure 3.4. The results of genome scan of AS-30CQxAJ cross.

Untreated (a) and treated with 5 (b), 10 (c) and 20 (d) mg of CQ $\text{kg}^{-1} \text{day}^{-1}$. The results of pyrosequencing assays are plotted against the cumulative genome distance described in section 3.3.8.3. The lines are separating the different chromosomes. The results are presented either as the proportion of sensitive strain (untreated population) or as comparative index (treated population).

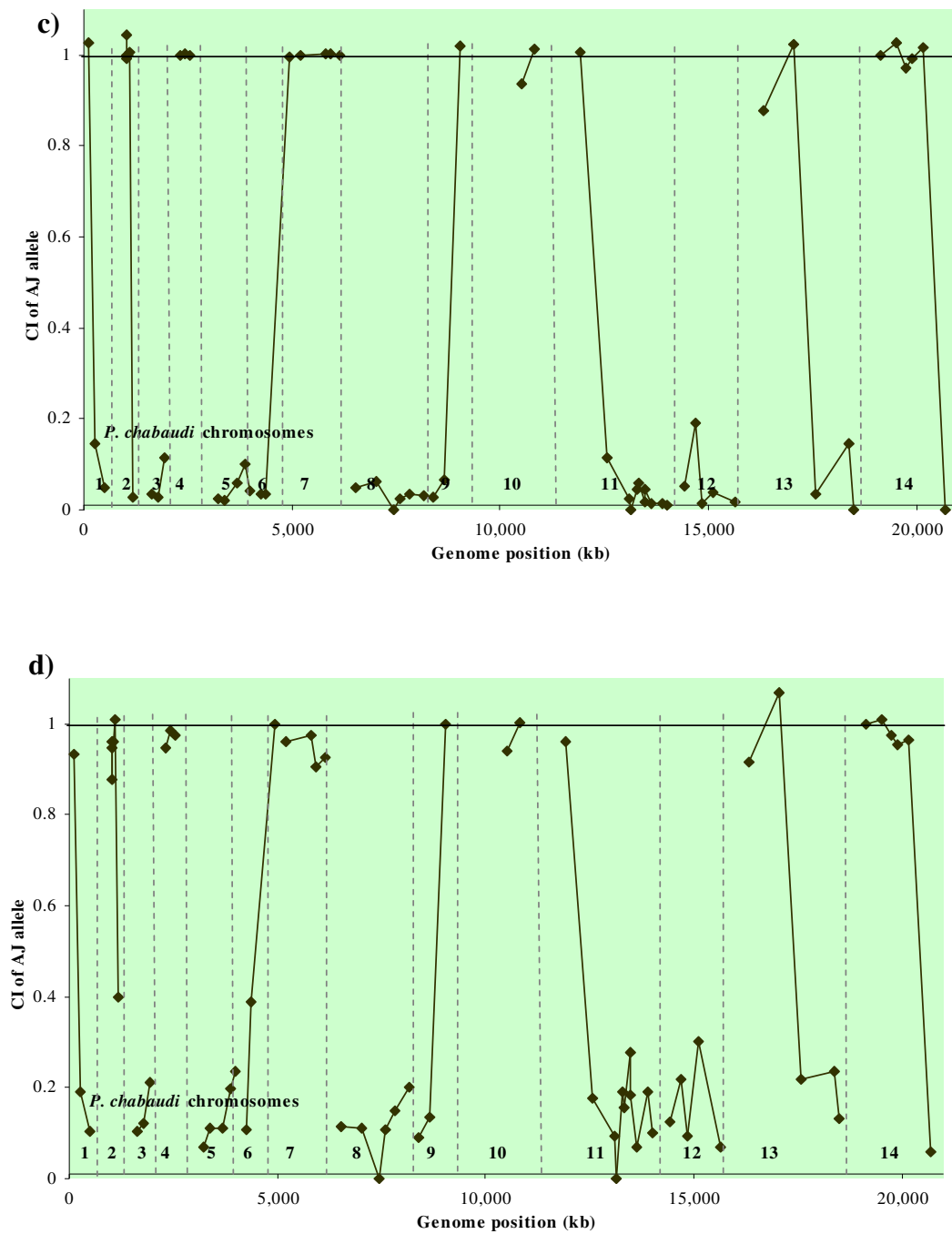


Figure 3.4 The results of genome scan of AS-30CQxAJ cross.

Untreated (a) and treated with 5 (b), 10 (c) and 20 (d) mg of CQ kg⁻¹ day⁻¹. The results of pyrosequencing assays are plotted against the cumulative genome distance described in section 3.3.8.3. The lines are separating the different chromosomes. The results are presented either as the proportion of sensitive strain (untreated population) or as comparative index (treated population).

3.5. Discussion

3.5.1. The AS-30CQxAJ cross

Several data suggest that a cross between the resistant and sensitive parent was successfully generated. A number of mosquitoes with ≥ 10 oocysts were sampled, suggesting quite intense transmission. Moreover, after propagation of the progeny, proportional sequencing indicated the presence of both parental alleles of *pcdhps* gene (chr 14) in the parasite population, in very even proportions (around 60:40). Although this single locus may not be representative for the rest of the genome, the fact that both parasites are represented in proportions very similar to those in the original inoculum (50:50) supports the speculation that no major allele loss/selection was observed in mosquito transmission. Also the behaviour of the cross during selection (very similar to those of mixtures of resistant and sensitive parents in the same conditions) suggests a balanced content of resistant and sensitive parasites.

An essential factor decisive for cross quality however is how many independent recombinant parasite genotypes are present in the progeny: The amount of recombination events is decisive for quality of the cross and the mapping of traits.

In total the progeny of ~1200 oocysts were recovered. Therefore, assuming the random mating and equal proportions of both strains in the pool of gametocytes contributing to the fertilisation, around 600 resistant recombinant clones were generated, which should be enough to map the major resistance loci. Unfortunately, there is no easy way to estimate how many recombinants are present in the population so these numbers could not be confirmed.

3.5.2. The untreated population – the selection for multiple loci for virulence?

In the untreated population a uniform, high percentage of AJ alleles at all the tested loci was observed. A possible explanation would be that this high baseline reflects the high proportion of AJ gametocytes in the mosquito meal/better survival of AJ genotypes in mosquito host etc. In that case the population composed from 80% of AJ and 20% of AS gametes would manage to recombine and produce progeny). However the proportion of alleles in samples taken soon after propagation suggested more equal proportions of both strain. Therefore, it was concluded that

parasites with higher-than-average combinations of AJ alleles were selected during growth in untreated mice.

Because in mixed infection without the drug treatment, AJ grows faster than AS-30CQ (see section 2.5.4.) selection of AJ alleles at loci underlying this “growth advantage” was expected. This trait is likely to be multigenic, involving genes expressed in many different pathways like DNA replication, erythrocyte invasion and re-modelling or immune system evasion etc. All of these processes may undergo some level of selection during the rounds of asexual growth. If the number of selected loci is high enough (*e.g.* 2-3 on every chromosome), instead of few isolated “selection peaks” we will observe a uniform level of selection across the whole genome. In other words, the parasites that inherited a significant portion of AJ genome will grow better than the ones that are carrying mostly AS alleles, effectively regardless of their position on the genome. Moreover our population contains a significant amount (~25% if the assumptions mentioned in section 3.5.1. are fulfilled) of AJ parasites being the results of mating of two AJ gametes. Presence of the fast growing AJ parasites that were likely to constitute a substantial proportion of infection was likely to uniformly increase the percentage of AJ allele at all loci, thus further increasing the baseline of genome scan (even for non-selected chromosomes). It is most likely explanation of observed results.

3.5.3. The treated population – many loci under selection or sampling error?

In contrast, in all three treated populations, a highly bimodal distribution of allele frequencies was observed (most markers showed either ~0% or ~100% AJ proportion). In the two highest drug concentrations almost no markers with intermediate allele frequency were observed. Moreover, large fragments of the genome (for example across adjacent markers) appear to have been inherited from one or other of the same parent.

One explanation would be that these selection signatures are the biological reality. The chloroquine resistance might be controlled by say 10 loci (which would be in a striking contrast with low CQ resistance in AS-3CQ mapped to a one single locus (Hunt et al., 2004b)) and the strong selection on them would decrease the

frequency of the whole group of surrounding markers. The rest of the genome would be selected for AJ alleles (rapid growth) as proposed in the previous section. This might result in the observed polarisation. In that case, the treatment would be selecting for one specific combination of parental genotypes – a “superparasite” being the perfect balance between the high drug resistance and high growth rate.

There is however another, more worrying possibility. Similar results could be obtained if a very low number of recombinants was sampled in the quantitative genotyping. Whole chromosomes would appear to be selected because few of the resistant recombinants happened to carry it. It would also explain the sudden changes of frequency between the neighbouring markers (what we call the scatter of the data) – it would simply reflect the single crossing over event in one major recombinant. In that case instead of optimal genotype(s) the scan would represent an extremely small number of recombinants that by chance managed to dominate the infection.

3.5.4. The loss of recombinants?

The main concern arising from the genome scan of the treated populations is that a very low number of recombinants may be present in the analysed progeny either because a very low number were generated during the cross or because of massive loss that occurred during propagation and passage in the generated population. I found the latter more likely. Unfortunately, maintaining *Plasmodium* parasites in the laboratory involves bottlenecks where parasite numbers are greatly reduced; for example, passage from mosquito to mouse, from one animal to another, freezing down and rethawing. Bottlenecks will eliminate some of the parasites and their genotypes. Moreover, each day of an infection gives parasite the chance to compete with each other. In the present case, previously obtained data suggests that there are fitness differences between the resistant and sensitive strain (described in chapter 2) and possibly fitness costs of the chloroquine resistance mutation itself (see section 2.5.5. of Chapter 2). Thus most of the parasites eliminated over time may tend to be resistant recombinants.

In that situation the original plan to increase the number of pyrosequencing markers in order to obtain better resolution of putative selection valleys was

discarded as it was unlikely to substantially improve the quality of the data. Instead I focused my efforts on generating better quality genetic material as described in chapter 4.

4. Generation and genetic analysis of AS-30CQ x AJ backcross

4.1. *Introduction*

This chapter features the backcross of the progeny of AS-30CQ x AJ cross to the sensitive parent AJ and its analysis in the search for loci involved in the resistance to chloroquine and artemisinin. These new experiments exploit previous work and introduce some modifications to the original protocol, aimed to improve the quality of the cross, selections and data analysis.

4.1.1. Rationale for backcross

In Chapter 3, I described the cross between AS-30CQ strain (selected for high level of chloroquine resistance and spontaneously artemisinin resistant, see section 1.7.2.) and a genetically different sensitive strain, AJ. The conclusion reached was that the quality of the original cross was not sufficiently diverse to study the resistance loci in sufficient detail – either because the trait is determined by multiple (>10) loci or because the tested population is dominated by a small number of random recombinants.

In such a case the backcross to the sensitive parent can be used to improve the results. It is preferred to the generation of a new cross as, in addition to creating new recombinants, it introduces additional crossing-overs in already existing ones and saturates the population with alleles from sensitive parent. It improves the genome scan in at least two ways (Fig. 4.1). Firstly the increased percentage of sensitive alleles at all the loci lifting the baseline of the genome scan and eliminate the majority of false positives that may appear as selection artefacts. Therefore “real” selection valleys are easier to identify. Secondly, the increased amount of recombination events narrows the selection valleys, allowing its lowest part to be mapped with higher resolution.

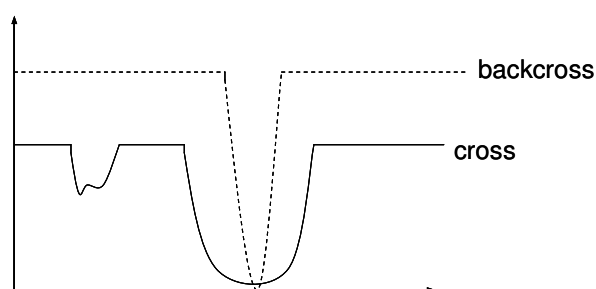


Figure 4.1 The improvement of data quality by the backcross.

The full line presents the hypothetical genome scan of the cross and the dashed one of the subsequent backcross. The valley on the left side of the diagram (in the cross) represents the false positive – an accidentally selected part of the genome, eliminated in the backcross, and the right one- a selection valley selected by the resistance.

4.1.2. Modifications of LGS protocol- aiming for a higher number of recombinants

As mentioned previously, one of the main concerns arising from the analysis of original cross is that only a small handful of resistant recombinants may be present in the final population. Therefore the main objective of the backcross should be to obtain high numbers of progeny and make sure that they survive to the selection.

A number of modifications designed to generate and maintain higher numbers of recombinants were considered:

- The donor mice for mosquitoes' infections were treated with erythropoietin. This drug is known to increase the numbers of gametocytes in animals by stimulating the production of reticulocytes (parasites infecting reticulocytes are more likely to become gametocytes) (Reece et al., 2005). Increased gametocytaemia before the feeds should result in increased transmission to the midguts of mosquitoes and in consequence generate more recombinants
- At various stages parasites were grown independently in untreated and treated mice and then pooled. This would retain the sensitive, fit parasites (that are going to be transmitted in untreated infection) and the highly resistant, slower growing parasites that can survive in treated infections, in order to supply a diverse population of recombinants for cross or drug-selections.
- Laboratory procedures (passages, deep freezing) that may cause 'bottlenecking' prior to selection experiments were minimised.

4.1.3. Modifications of selection procedure – the use of low, intermediate and high drug concentration

In Chapter 3 selection protocols aimed to select the highly resistant recombinants and thus map onto one genome scan all of the loci involved (see section 3.1.2 in Chapter 3). The results suggested that a large number of loci may be selected. In this chapter, the backcross is treated with multiple drug concentrations that were designed to select recombinants of various level of resistance and thus reveal a fine resolution of loci conferring low resistant, intermediate and highly resistant phenotypes. These multiple selections can be used to maximise the chances to map all of the responsible loci and the relationship between them. For example, if the loci are acting independently, we are likely to select all of them with low drug concentrations although none of them will be fixed in the population. The higher doses would gradually deepen the valleys (Fig. 4.2a). On the other hand, if some loci contribute only to increase a pre-existing resistance but do not confer any advantages by themselves, with increasing drug concentration we may observe first the selection on one locus then at two etc (Fig 4.2 b).

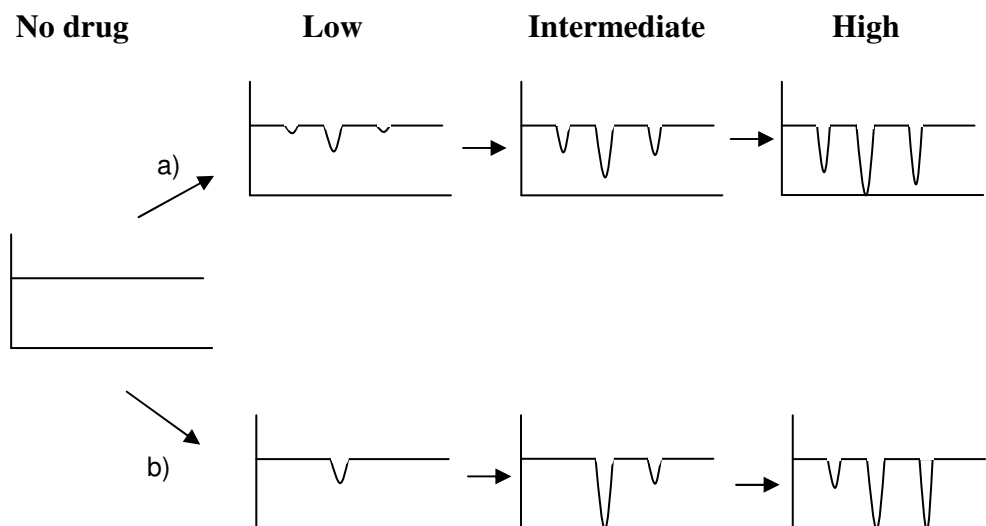


Figure 4.2 Predicted patterns of appearance of selection valleys in an experimental population. All the graphs are schematic representations of whole genome scan. The two paths are representing the loci either acting independently and selected together (a), or acting one on a top of another and selected in that order (b)

4.2. *Experimental design*

Figures 4.3 and 4.4 present a schematic of backcross and selections of obtained progeny.

4.2.1. Pre-selection of the cross and genetic analysis of obtained material

In order to generate the mixture of parasites containing both resistant and sensitive recombinants from the previous cross, the parasites were recovered from liquid nitrogen and pre-selected with 0, 1.5 and 10 mg of CQ kg⁻¹ day⁻¹ for 3 days (Fig 4.3).

After being used as donors for the backcross, the mice from three pre-selected groups were bled and parasite DNA was isolated for genetic analysis. The genome scan was performed on each sample as described in section 3.2.4. using 59 pyrosequencing markers (see Appendix C for details). This was primarily done in order to estimate the quality of the material introduced into the backcross (the proportional sequencing of one locus could no longer be used for this purpose because of the presence of recombinants in the mixture). However this quality assessment also supplied a new set of genome scans, that could be used for the identification of the resistant loci (if the data quality is better than the one obtained in the initial selection in Chapter 3).

4.2.2. Generation of the backcross

The populations generated in the pre-selection were mixed with sensitive parent AJ (using the protocol described in section 2.3.2.3.) in following proportions:

AJ:	30%
cross treated with 0 CQ	10%
cross treated with 1.5 mg of CQ kg ⁻¹ day ⁻¹	30%
cross treated with 10 mg of CQ kg ⁻¹ day ⁻¹	30%
<hr/>	
	100%

The prepared mixture was injected into the set of five C57 mice pre-treated with erythropoietin and grown for 5 days before mosquito feeds. The cross protocol described in chapter 3 was followed.

The progeny were recovered into a group of 10 animals, with sporozoites from each mosquito cage injected into 2 specific mice. The resulting parasites were subjected to two series of selections, called Selection 1 and Selection 2 (Fig 4.3 and 4.4, respectively). All CQ concentrations used in these selections were established in Chapter 2. The ART concentrations were determined from the similar phenotyping experiments performed by Richard Fawcett and Louise Rodrigues (data not shown).

4.2.3. Selection 1

Selection 1 aimed to discover the contribution of loci to the different levels of CQ resistance. It was performed immediately after the progeny was recovered from the backcross (the use of the larger numbers of mice made it possible to skip the amplification step used in the original cross procedure). 4 groups containing 5 mice each were set up and treated for the first 3 days of infection as follows:

- **no drug** –allows the growth of all the parasites and can be used as a control for other groups
- **1.5 mg CQ kg⁻¹ day⁻¹** –a low concentration that should separate of all the resistant parasites (even those with very low level of resistance) from the sensitive ones. It has to be reminded however that the previous results obtained with that concentration were very variable (see Fig 2.8) and some recrudescence of sensitive parasites was possible in with that treatment regime.
- **5 mg CQ kg⁻¹ day⁻¹**– the intermediate concentration that had been proven to kill the sensitive parasites and significantly delay the growth of low resistant ones. We aimed to select the intermediate and highly resistant progeny with this treatment.
- **10 mg CQ kg⁻¹ day⁻¹**– the high concentration chosen to select for the most resistant parasites only.

4.2.4. Selection 2

Selection 2 was subsequently designed to optimise the data obtained in Selection 1. The starting parasites were obtained from a sample of parasites deep-frozen from the mice used to recover blood forms from sporozoites (Fig. 4.4). These parasites were pre-selected at 0 and 10 mg CQ kg⁻¹ day⁻¹ and then pooled. This pool was selected in 7 group of 4 mice each, at 0, 1.5, 3, 10 and 20 mg CQ kg⁻¹ day⁻¹.and 100, 200 mg ART kg⁻¹ day⁻¹.

4.2.5. The genetic analysis of backcross progeny

The parasites were harvested before the peak of infection and analysed using the set of quantitative pyrosequencing markers as described in previous chapter (see section 3.2.4.). For the selection 1, 60 markers were tested before it was judged that the addition of supplementary ones will not result in substantial data improvement (see section 4.5.1.). For selection 2 all available markers (and some new ones designed in the loci of interest) were tested in order to map the potential loci involved in the resistance with the high resolution. In total the data from 93 assays were obtained (see Appendix C). Additionally the data from selection 2 was subjected to statistical analysis in order to assign the statistical value to the putative selection valleys.

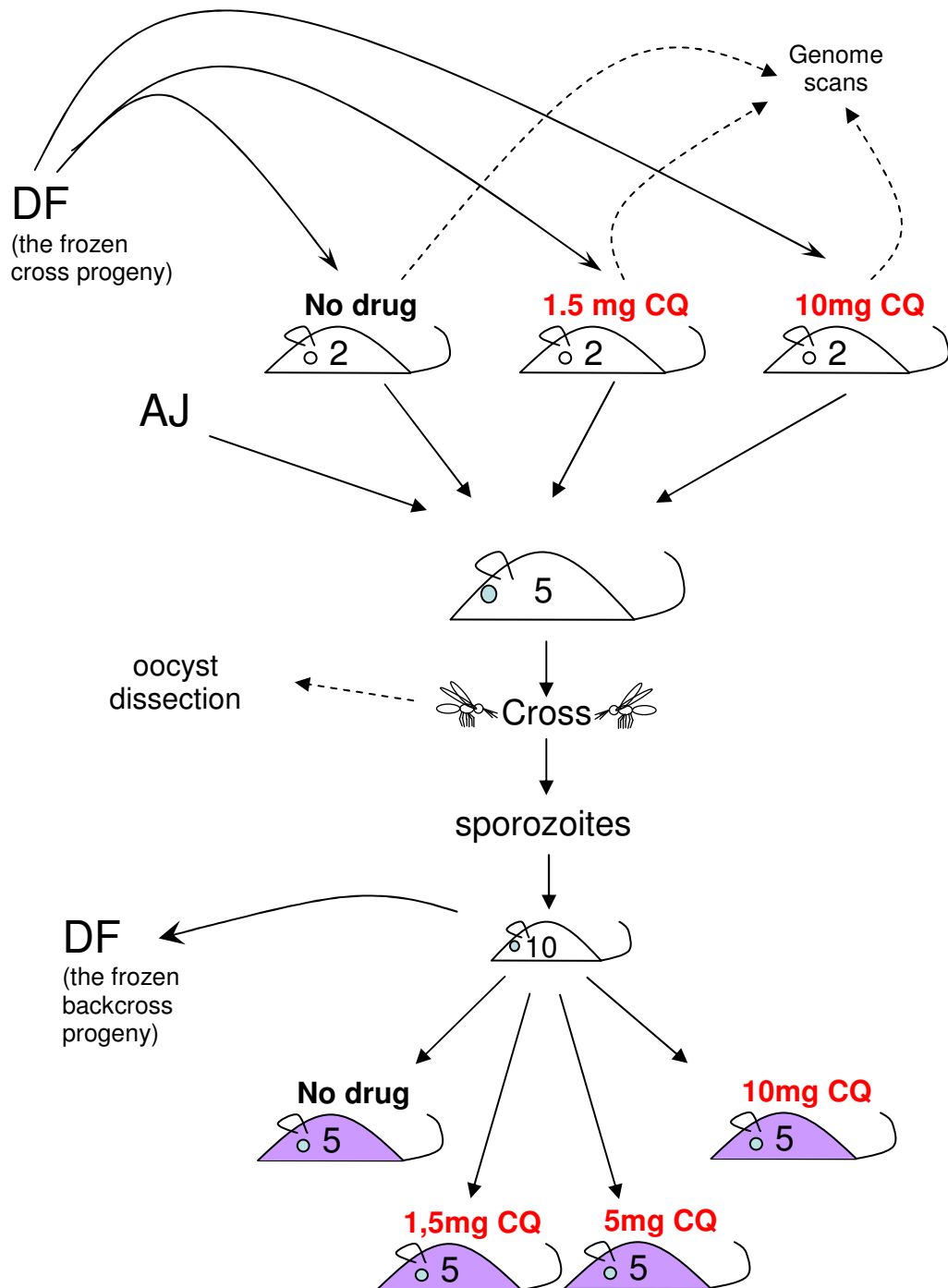


Figure 4.3 The backcross and selection 1 of its progeny.

Parasites from the AS30CQ x AJ cross were recovered and pre-selected with 0, 1.5 and 10 mg CQ/kg/day. The surviving parasites were pooled and mixed with AJ parasites to generate the backcross. The resulting progeny were recovered in 10 mice and both frozen down (for use in Selection 2) and selected in Selection 1. The solid arrows represent the transfer of the parasites; the dashed arrows represent 'monitoring' steps. The number inside each mouse represents the number of animals used in that step. The drugging regime used is shown above each mouse (all animals were treated on days 0-2). DF-parasites put into/taken from deep freeze; violet shading - mice used for LGS genome scans.

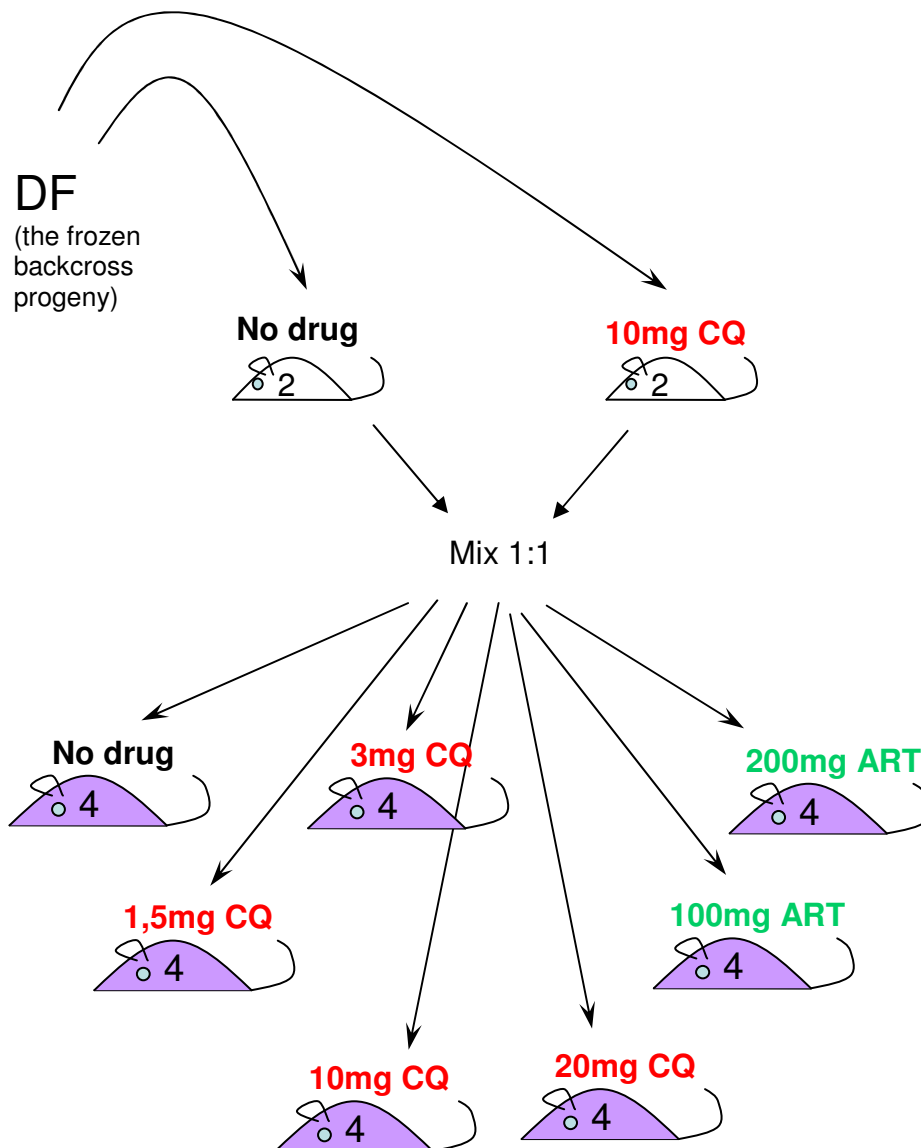


Figure 4.4 Selection 2 of the progeny of the backcross

The parasites from backcross recovery were taken from liquid nitrogen and pre-treated with 0 or 10 mg of CQkg⁻¹ day⁻¹. The mix of both populations was used to inoculate mice for selection. The full arrows represent the transfer of the parasites. The number inside each mouse represents the number of animals used in that step. The drugging regime used is shown above each mouse (all animals were treated on days 0-2). DF - parasites taken out of deep freeze; violet shading - mice used for LGS genome scans.

4.3. Materials and Methods

Most of the Materials and Methods used in this chapter were described in Chapter 3. Only the details of additional methods are mentioned here.

4.3.1. EPO treatment

The C57 mice used for mosquito feeds were treated with mouse erythropoietin (Roche) for 5 consecutive days starting two days before the parasite injections (d -2 to 2 p.i.). The EPO was diluted with water to 100U l^{-1} and injected intraperitoneally in a volume of 100 μl per mouse per day.

4.3.2. The recovery of the cross progeny from liquid nitrogen

The cryopreserved parasites were defrosted in ice and the deep-freeze solution was removed by a series of three washes in 1) 12% sodium chloride. 2) 1.6% sodium chloride 3) 0.9% sodium chloride, 0.2% glucose. The recovered parasites were re-suspended in equal volume of citrate saline and injected intraperitoneally into uninfected mice in a volume of 100 μl per mouse

4.3.3. Artemisinin treatment

For artemisinin treatments, the pre-calculated amount of artemisinin powder (Amersham, UK) was weighed and dissolved in Dimethyl sulfoxide (DMSO) to obtain the highest needed concentration. Lower concentrations were achieved by further dilution with DMSO. The mice were force-fed with drug solutions using gavage needle.

4.3.4. Statistical analysis

The statistical analysis described in this section was designed and performed by Laurence Loewe, Dario Beraldi and Paul Hunt

In order to test whether observed selection valleys are significant, two different statistical approaches have been developed.

The first one is based on computer simulations. The population of virtual parasites (each containing 100 genetic markers spaced across the genome at distances equivalent to the ones used in the real experiments) was generated. The experimental design (generation of recombinants, rounds of asexual replications, selection of the population and sampling of genetic material for analysis) was modelled. Then, LGS scans were simulated under the null hypothesis (*i.e.* the hypothesis that we want to disprove with calculated significance) that the resistance is due to a large number of independently acting loci. With no dominant loci, no selection valley would be observed, so any fluctuations in the observed allelic frequency would be due to variation in allele sampling, experimental error etc. After running the simulations 500 times the frequency with which the selection valley of given depth appeared in the model could be calculated and used to assign the P-value of similar valley observed in the experimental data. For example, if a selection valley of a given depth would be observed in one out of 10 simulations the P-value of this valley when it is observed in our experiment is $P=0.1$ and therefore it is discarded as insignificant. If however we observe the selection valley that appears in one out of 10 000 simulations, with $P=0.001$ we consider it significant.

The second approach consists of comparing the mean of sliding windows containing 3-5 alleles from the same chromosome in the treated population, with the mean of all the alleles from untreated population. In order to make the test more stringent, the values of all AJ allele frequencies in the untreated population are reduced by an arbitrary percentage (25%), reflecting the overall decrease of frequency of AJ allele in treated populations (caused by the elimination of sensitive parental genotypes and part of the recombinants). To estimate the P-value, Mann-Whitney U-test was used because it does not require any prior knowledge about the data distribution. With this method the *groups* of markers that are decreased more than average can be identified.

4.4. Results

4.4.1. Reselection of cross progeny

The reselection of progeny recovered at 0, 1.5 and 10 mg of CQ kg⁻¹ day⁻¹ took place prior to the genetic cross. All animals developed parasitaemia as expected (data not shown). After the collection of blood necessary for the backcross, the genome scans were performed on the remaining material in order to assess the quality of the populations introduced in the mixed infection (Fig 4.5).

The presence of alleles from both parents was observed throughout the genome in all the three populations. In general the data replicated the results from the selection experiments in Chapter 3 with the lack of treatment increasing the proportion of AJ allele and 10 mg of CQ kg⁻¹ day⁻¹ causing the high representation AS alleles in many loci. Additionally the variation was observed between the markers suggesting the presence of recombinants in all infections.

Therefore the mixture of the pre-selected populations with AJ parent was likely to result in a good backcross inoculum, containing a variety of resistant and sensitive recombinants.

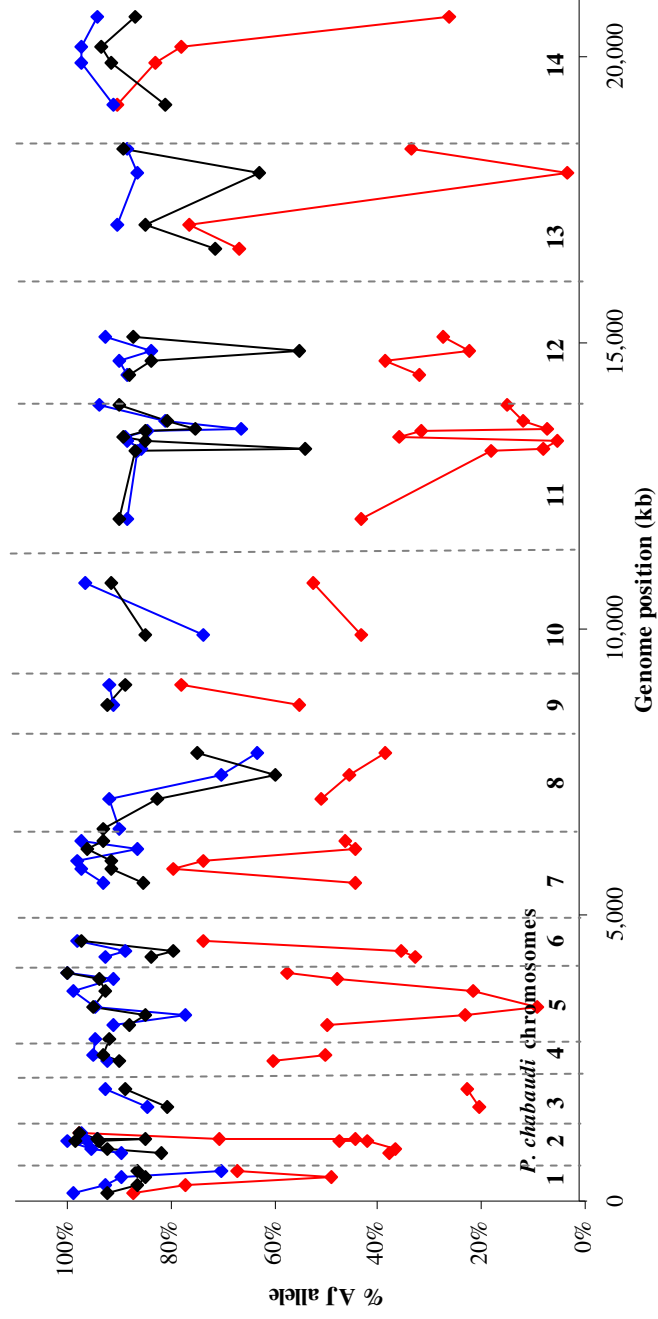


Figure 4.5 The results of genome-scans performed on preselected populations. The frequency of a few genetic markers on each chromosome was tested in order to ensure the presence of alleles of both parents in the material for the backcross.

4.4.2. The generation of the backcross

The mixture of pre-treated infections and AJ parasites was set up as described in section 4.2.1 and fed to 5 mosquito cages (250 mosquitoes each). The results of oocyst dissections are presented in Table 4.1 together with an estimation of oocysts present in each of the cages. This analysis suggests that a total of ~8800 oocysts were produced. All the female mosquitoes were dissected and the sporozoites recovered in new group of 10 mice.

Mosquito number	Cage 1	Cage2	Cage 3	Cage4	Cage5
1	12	24	~60	6	1
2	0	38	20	24	3
3	6	~30	1	9	5
4	~30	1	17	12	0
5	27	5	~40	21	9
6	0	25	26	0	2
7	10	~50	15	7	0
8	1	17	~30	~35	0
9	0	22	3	0	2
10	3	~40	2	2	0
Average	~8.9	~25.2	~21.4	~11.6	~2.2
Number of dissected mosquitoes	112	133	102	183	102
Estimation of oocyst in cage	996.8	3351.6	2182.8	2122.8	224.4

Table 4.1 Oocyst analysis during backcross generation.

The average number of oocysts per mosquito and number of dissected mosquitoes leads to estimates of the number of oocysts obtained from each cage.

4.4.3. The selection of the backcross

4.4.3.1 Selection I.

The parasites were grown in 4 groups of five mice treated with 0, 1.5, 5 and 10 mg CQ kg⁻¹ day⁻¹

All parasites developed parasitaemias (Figure 4.6). In the untreated group parasites appeared on day 3. In the treated groups the onset of infection was delayed in proportion to the concentration used (parasites emerged on day 5, 10 and 12 respectively). Groups of mice were killed on day 6, 8, 13 and 16 respectively and DNA recovered for genetic analysis.

4.4.3.2. Selection II.

The parasites were selected in 7 groups of mice – untreated and treated with 1.5, 3, 10 or 20 mg CQ kg⁻¹ day⁻¹ or with 100 and 200 mg ART kg⁻¹ day⁻¹. The growth of parasites is presented in Figure 4.7. The parasites grown in CQ-treated and untreated animals behaved as in the previous experiments; *i.e.* that the delay of parasitaemia was proportional to the drug concentration used. For example in the group treated with 3 mg CQ kg⁻¹ day⁻¹ parasites appeared on day 7 and reached ~25% on day 10 when mice were bled for parasites. Artemisinin treatments also delayed the appearance of parasites and the growth in these infections was comparable to those observed with the highest CQ treatments. For example in the group treated with 100 mg ART kg⁻¹ day⁻¹ parasites were growing in parallel with those treated with 10 mg CQ kg⁻¹ day⁻¹ and those treated with 200 mg ART kg⁻¹ day⁻¹ were similar to those treated with 20 mg CQ kg⁻¹ day⁻¹.

Parasites were harvested on the following days – untreated and those treated with 1.5 mg CQ kg⁻¹ day⁻¹ of CQ on day 7, those treated with 3 mg CQ kg⁻¹ day⁻¹ on day 10, those treated with 10 mg CQ kg⁻¹ day⁻¹ or 100 mg ART kg⁻¹ day⁻¹ on day 11 and those treated with 20 mg CQ kg⁻¹ day⁻¹ or 100 mg ART kg⁻¹ day⁻¹ on day 13.

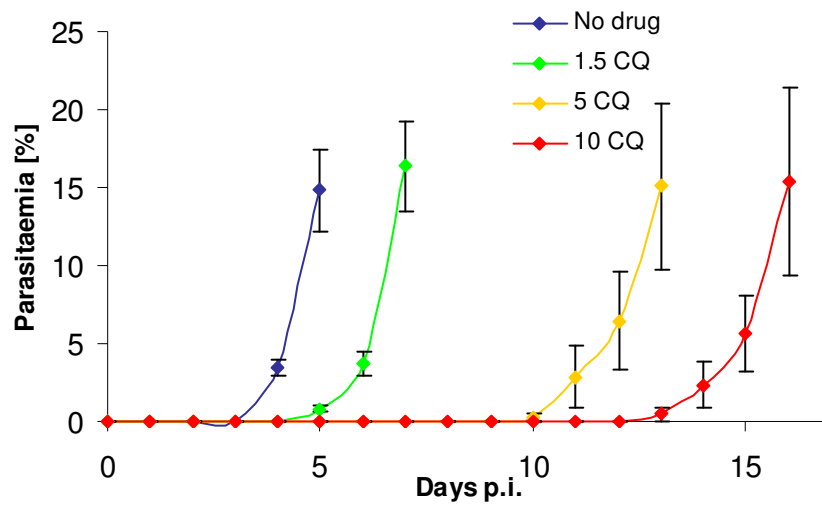


Figure 4.6 The growth of AS30CQ x AJ backcross during the selection I

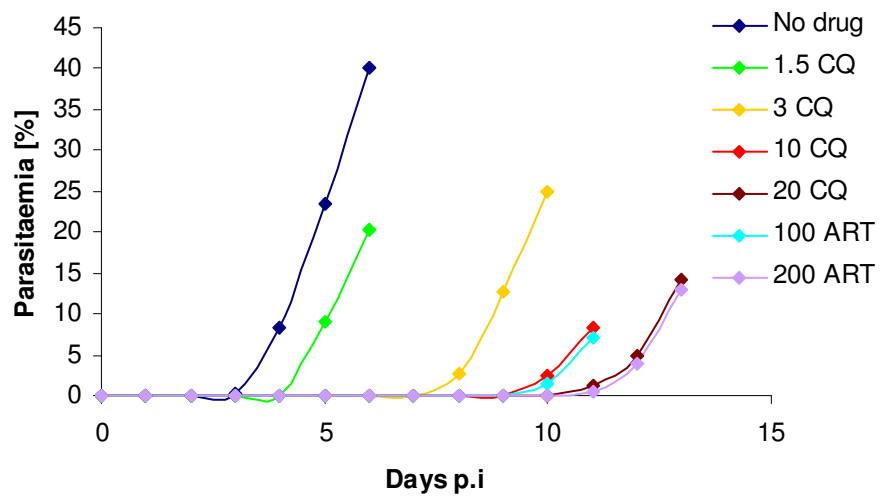


Figure 4.7 The growth of AS-30CQ x AJ backcross during selection II.

4.4.4. The genetic analysis of the backcross – selection I

The frequencies of alleles at 60 parental markers were tested in all selected populations. The data obtained (AJ allele percentages) is presented in Figure 4.8a-d. Comparative indices (allele proportions in drugged samples relative untreated samples) are shown in Fig 4.9

4.4.4.1. The untreated population

Here, a uniformly high ($90 \pm 1.1\%$) percentage of AJ markers is observed along the genome (Fig. 4.8 a). It is consistent with the behaviours of the original AS-30CQ x AJ cross that showed similarly high proportions of AJ markers in untreated infections

4.4.4.2. The 1.5 mg CQ kg⁻¹ day⁻¹ treated population

Here also the majority of the markers showed a high proportion of the sensitive allele ($80.1 \pm 2.2\%$) similar to the untreated population (Fig 4.8b). This time however one strongly reduced group of markers can be observed on chromosome 11. The lowest marker on this chromosome (pcpf 06-1338) is reduced to 23.3% (compared to 72.2% in untreated population).

4.4.4.3. The 3 mg CQ kg⁻¹ day⁻¹ treated population.

The baseline AJ allele frequency is $72.1 \pm 2.5\%$ and the markers are much more scattered. (Fig 4.8c). Again the most prominent feature of the scan is the selection on chromosome 11 (now greatly deepened - the lowest marker, pcpf 06-1004, is 18.6 %). Moreover, some other markers are reduced in comparison to both untreated and treated with low CQ dose ($1.5 \text{ mg CQ kg}^{-1} \text{ day}^{-1}$). Notably, markers on chr 3 shows signs of selection -the three markers on this chromosome (pcpf 02-0275, pcpf 02-0452, pcpf02675) are reduced from 90, 90, 82 % (untreated) to 49, 54 and 58% respectively. Other supplementary loci may possibly be present on chromosomes 5, 6 or 12.

4.4.4.4. The 10 mg CQ kg⁻¹ day⁻¹ treated population

The 'baseline' is further reduced with a mean AJ allele frequency of $70 \pm 2.6\%$ (Fig 4.8d). The scatter of the markers is increased in this selection, making the interpretation of a selection valley more difficult. Markers on chr 11 and chr 3 are still the lowest in the genome. However, even at this relatively high concentration of drug, none of the AS alleles on chromosome 11 are selected completely. Even the lowest point of the selection valley (marker pcpf 06-1154) is 38.9%, showing that a surprisingly high proportion of the parasites which apparently carry the sensitive allele. Moreover the selection valley appears to be more shallow than that observed with 3 mg CQ kg⁻¹ day⁻¹.

The selection valley on chr 3 is slightly deeper than in the previous group, but none of the markers are completely reduced.

Additionally the markers on many other chromosomes are decreased in comparison with the previous concentration (see for example chr 4, 5 and 6) however it is difficult to establish whether they represent minor selection signatures or the general decrease of frequency across the genome.

An interesting feature can be observed on chromosome 2. This chromosome is characterised by a high percentage of AJ alleles in all previous selections (the mean marker density of 93.2%, 92.3% and 84.7% for untreated, 1.5mg and 3mg groups respectively). However in 10 mg of CQ kg⁻¹ day⁻¹ group, the % of all AJ markers on this chromosome is on average 64.3 %. It is the most visible difference between this scan and the previous one.

It is therefore possible that an additional locus involved in CQ resistance is located on this chromosome.

4.4.4.5. Summary of selection 1

In summary, all three chloroquine treatments selected for AS markers situated on chromosome 11. However in each case different pyrosequencing marker is the lowest point on this chromosome. Also the shape of selection valley changes with each drug concentration. The intermediate and high treatment seems to select at least one additional locus on chromosome 3 and possibly, at the highest CQ concentration, one on chromosome 2. However none of the concentrations reduces AJ alleles of

any of the markers to zero which would be expected if at least one locus required for resistance exists.

The interpretation of these data and the rationale for selection 2 is discussed in section 4.5.1.

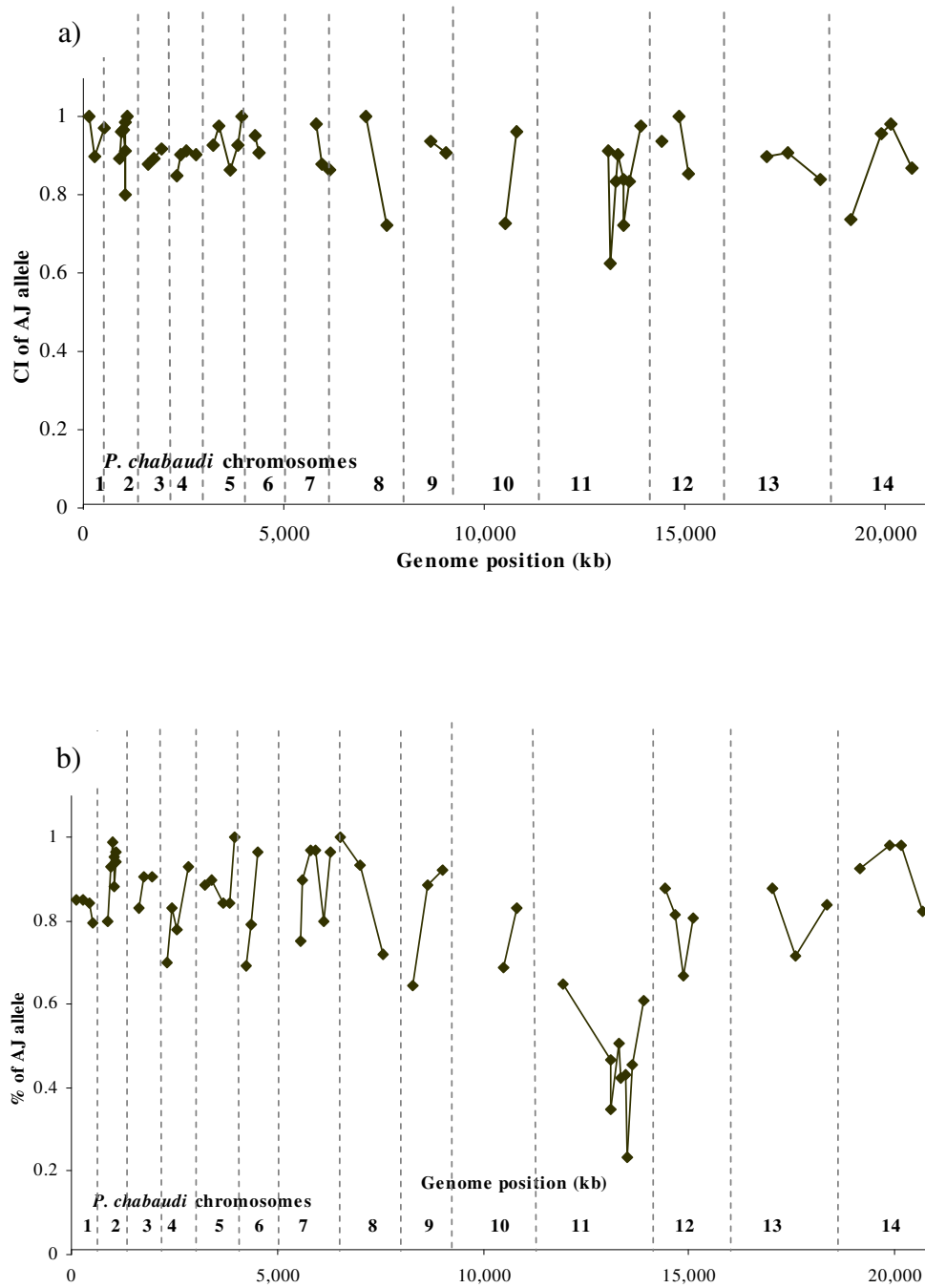


Figure 4.8 Whole genome scans (selection 1)

The graphs are representing the proportions of sensitive allele in the selected mixture: untreated (a) and treated with 1,5 (b), 5 (c) and 10(d) mg/kg/day of chloroquine

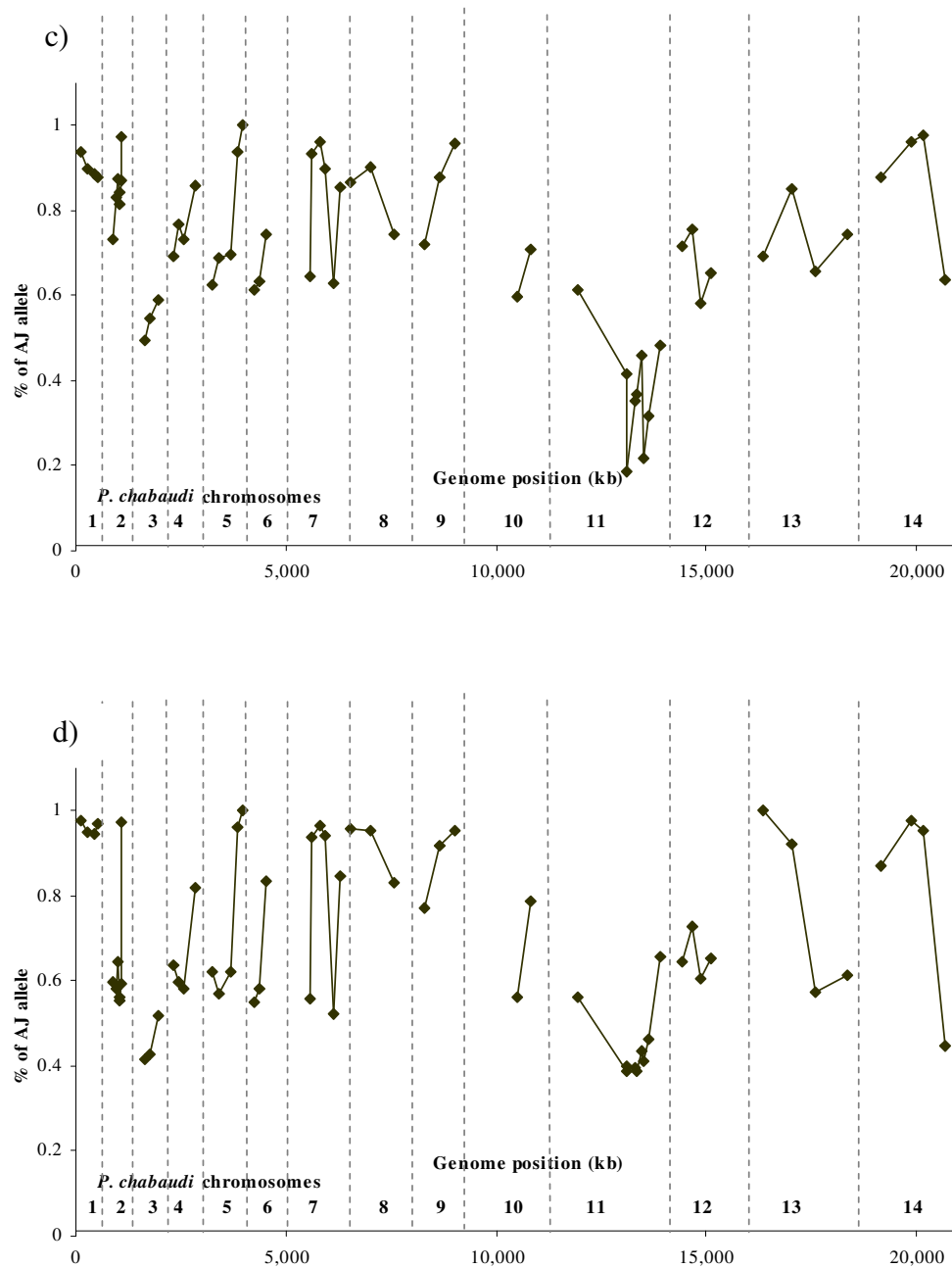


Figure 4.8.b Whole genome scans (selection 1)

The graphs are representing the proportions of sensitive allele in the selected mixture: untreated (a) and treated with 1,5 (b), 5 (c) and 10(d) mg/kg/day of chloroquine

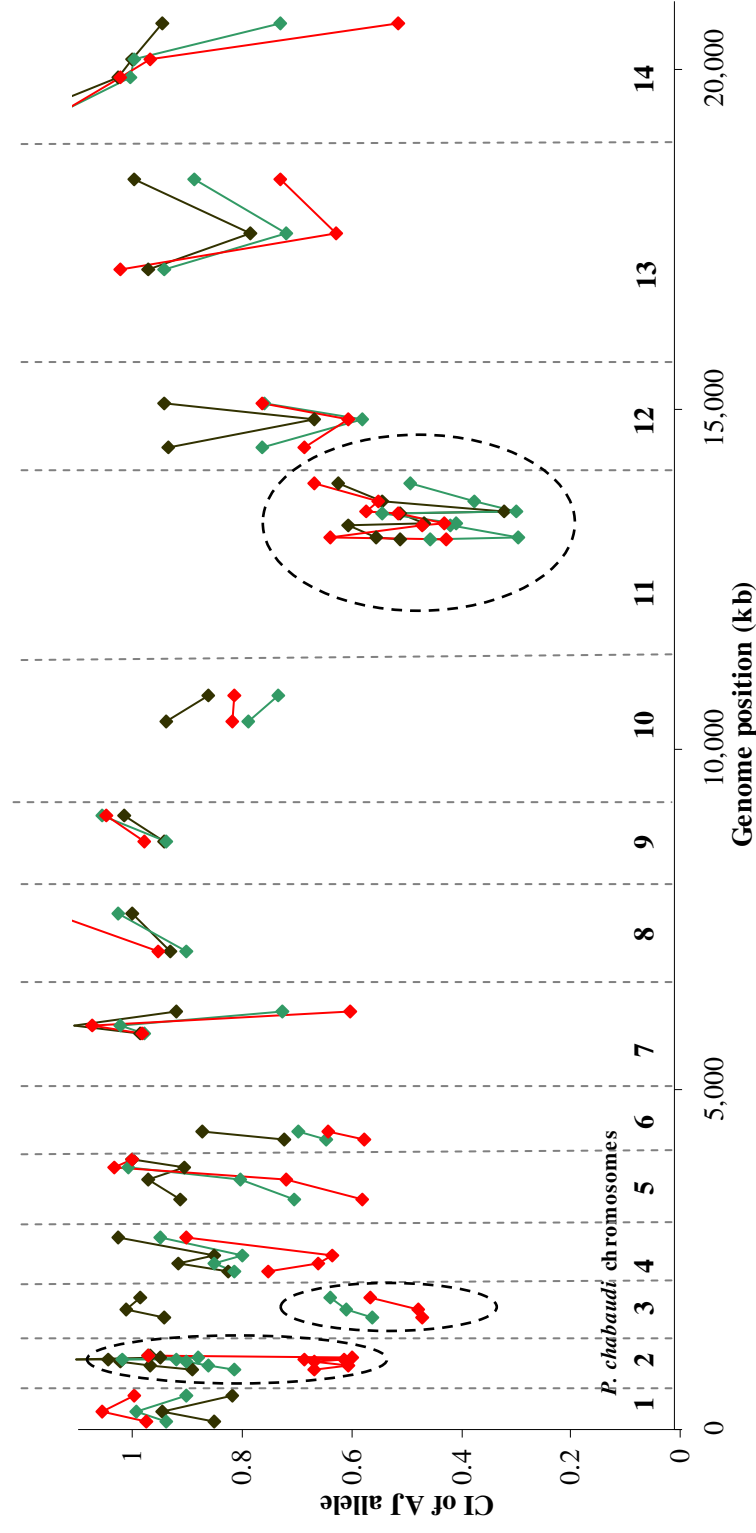


Figure 4.9 Selection 1 transformed data.
 ■ - 1.5 g/kg /day, ■ - 3mg/kg/day, ■ - 10 mg/kg/day. The circles are indicating loci of special interest.

4.4.5. The genetic analysis of the backcross – selection 2

Here, the frequencies of AJ alleles at 93 genome-wide markers were determined using the library of pyrosequencing assays (Fig 4.10 – 4.16).

4.4.5.1. The untreated population

The mean percentage of AJ alleles genome-wide was ($87.6 \pm 0.8\%$) (Fig 4.10). This compares with 90% in Selection 1. No selection valley was observed.

4.4.5.2. The population treated with $1.5 \text{ mg CQ kg}^{-1} \text{ day}^{-1}$

The mean percentage of AJ alleles genome-wide was $87.7 \pm 0.7\%$ (Fig. 4.11). No selection valley can be observed. This was similar to the results observed in the untreated sample. Note, however, that this is in contrast with the data from selection 1 when a selection signature on chr 11 was observed with the same drugging regime.

4.4.5.3. The population treated with $3 \text{ mg CQ kg}^{-1} \text{ day}^{-1}$

Here, the mean percentage of AJ alleles genome-wide was $63.6 \pm 2.8\%$ (Fig. 4.12 a). This represents an apparent loss of AJ alleles across the genome, relative to those samples subject to lower doses of CQ. The most prominent features of this scan however are two clearly identifiable selection signatures.

The most dominant one is the clearly defined, reasonably symmetrical selection valley located on chromosome 11, between the markers pcpf06-1154 and pcpf12-2207 (Fig. 4.12 b). The marker showing the lowest percentage of AJ alleles is pcpf06-1338 with an AJ frequency of only 2.5 %. These data mean that there is almost complete selection for AS alleles within that locus. The statistical analysis revealed that this valley is highly significant (P value < 0.001 obtained with both statistical approaches taken).

A second region of selection was identified on chromosome 3 (Fig. 4.12 c). This time however the region of selection cannot be further narrowed – all three markers on this chromosome (pcpf 02-0452, pcpf02-0452 and pcpf02-0675) are inherited with very similar AJ allele frequencies (17.9, 17.0 and 23.3% respectively). Therefore a whole chromosome must be considered in the search for potential

resistant determinants. The statistical analysis showed that the valley is statistically significant (P -value = 0.034) with one of the exploited statistical methods (simulation based) but not significant with the Mann-Whitney test.

No other chromosome contains markers decreased below 30%.

These data confirm the results for Selection 1 where AS alleles of markers on chr 11 and chr 3 were also selected at 1.5, and 3 mg CQ/kg/day.

4.4.5.4. The population treated with 10 mg CQ kg⁻¹ day⁻¹

Here, a further decrease in AJ alleles is observed (mean percentage genome-wide equal to $52.5 \pm 3.5\%$ (Fig. 4.13 a). This time three potential selection signatures were observed – two previously identified (chr 11, chr 3) and another on chr 2. None of them is classified as significant with both used methods of statistical analysis (probably due to the higher noise level in AJ allele frequency compared to previous concentrations)

The valley situated on chromosome 11 (Fig. 4.13 b) was deeper and its bottom is defined by the pcpf06-1154 marker and reaches 1.5% (within the standard error of measurement) suggesting that all of the parasites in this population inherited this part of the genome from the resistant parent. Although its bottom was less clearly defined spanning from pcpf06-1001 to pcpf 13-2207 it was clear that the resistance is caused by the same locus that was observed with 3 mg of CQ kg⁻¹ day⁻¹.

A region of selection on chromosome 3 is also present (Fig. 4.13 c), but there appears to be no further selection relative to 3 mg CQ kg⁻¹ day⁻¹

A new, additional locus can be identified on chromosome 2 (Fig. 4.13 d), consistent with the data from selection 1 in which the AJ alleles on this chromosome were decreased in parasites treated with high but not with intermediate CQ concentration. The clear selection valley was identified between the markers pcpf01-265 and pcpf07-1006 with the lowest AJ marker (pcpf01-0158) reduced from 79.2% (in 3mg CQ kg⁻¹ day⁻¹ scan) to 14.0 %.

4.4.5.5. The population treated with 20 mg CQ kg⁻¹ day⁻¹

These data are very similar to those acquired for 10 mg CQ kg⁻¹ day⁻¹. The mean AJ allele proportion is only slightly lower ($50 \pm 3.7\%$) and no new loci under selection were discovered in comparison to the previous concentration (Fig 4.14 a).

The three selection valleys (on chr 11, 03 and 02) were still present (Fig 4.14.b-d). The lowest percentages of AJ alleles on these three chromosomes were 0.8 % (marker pcpf 06-1044 on chr 11), 13.4% (marker 02-0453 on chr 3) and 6.9% (marker 01-0158 on chr 2).

4.4.5.6. The population treated with 100 mg ART kg⁻¹ day⁻¹

The mean percentage of AJ allele along this genome scan is $60.5 \pm 2.8\%$ showing that ART treatment also selects for AS alleles in the population (Fig 4.15 a).

The main feature of the genome scan was one clear selection valley identified on chr 2, similar to that obtained with the highest chloroquine concentrations (P value <0.001 with both statistical methods used) (Fig 4.15 b). However, the selection was not complete. The proportion of the AJ allele of one marker only on this chromosome (pcpf01-0158) was decreased below ten percent (8.4%).

4.4.5.7. The population treated with 200 mg ART kg⁻¹ day⁻¹

Here, the mean percentage of AJ alleles along this genome was decreased to 54.5 ± 3.7 (Fig 4.16 a). Again the dominant selection valley was observed on chromosome 2 (Fig 4.16 b). The group of three markers on this chromosome (pcpf01-0197, pcpf01-0158 and pcpf01-0150) were reduced to below 10% and the lowest of them (pcpf01-0197) barely contained traces (>0.1%) of the AJ allele. Therefore at least part of this chromosome was under almost complete selection.

Interestingly AJ allele proportions on chr 11 are also strongly reduced, forming the selection valley similar to those generated with intermediate and high CQ treatments (Fig 4.16 c). For example, the proportion of the pcpf06-1154 (the lowest marker on this chromosome in this selection and also in CQ selections) was reduced to 3.8 % showing almost complete selection.

In addition one marker on chr 7 (pcpf04-0820) showed a very strong reduction. However as none of the others markers on this chromosome shown any signs of

selection (90% of AJ allele was observed in all of them). It was therefore possible that the observed result was an experimental error.

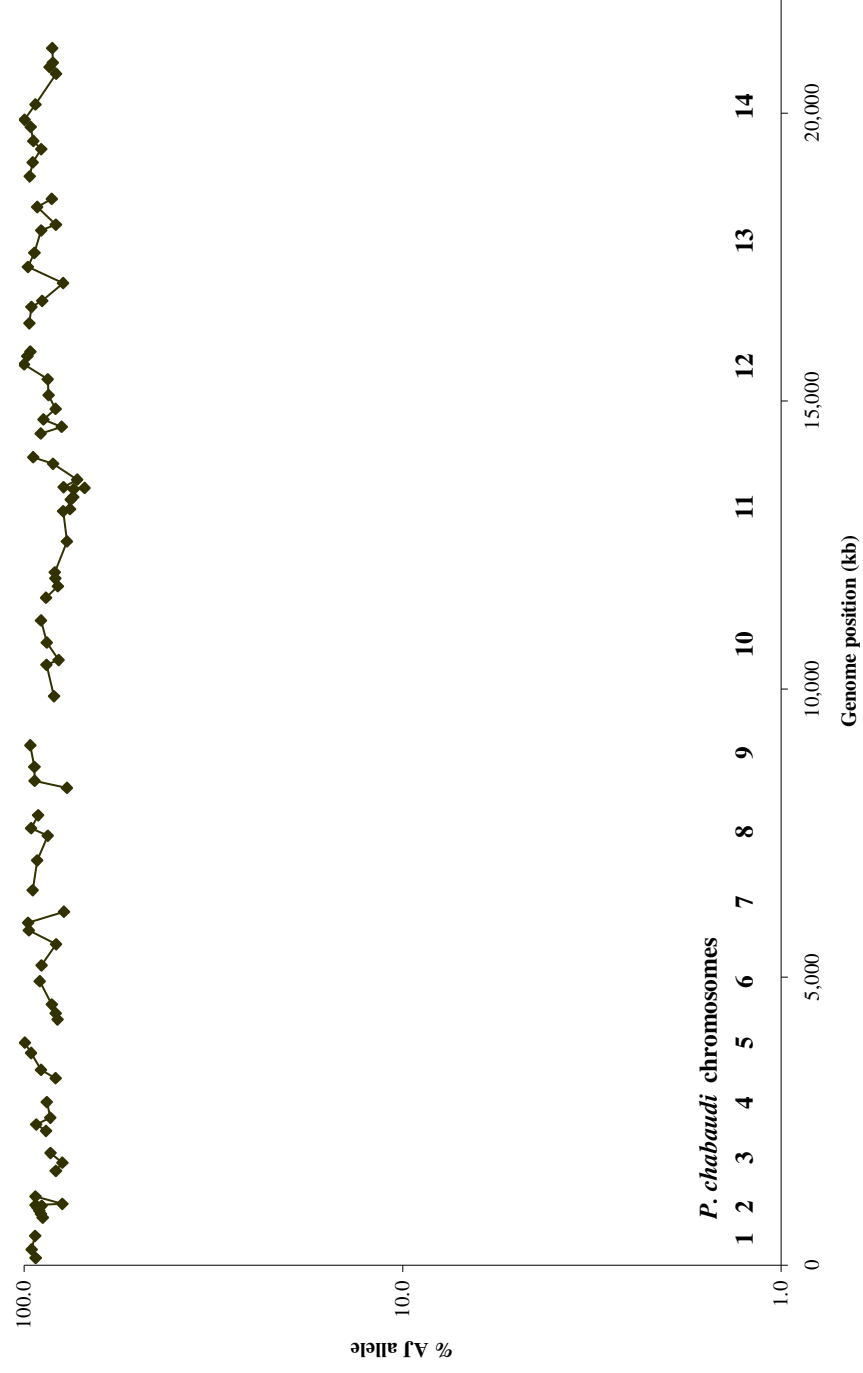


Figure 4.10 Selection 2 of AS-30CQ x AJ backcross - untreated
. Please note the logarithmic scale.

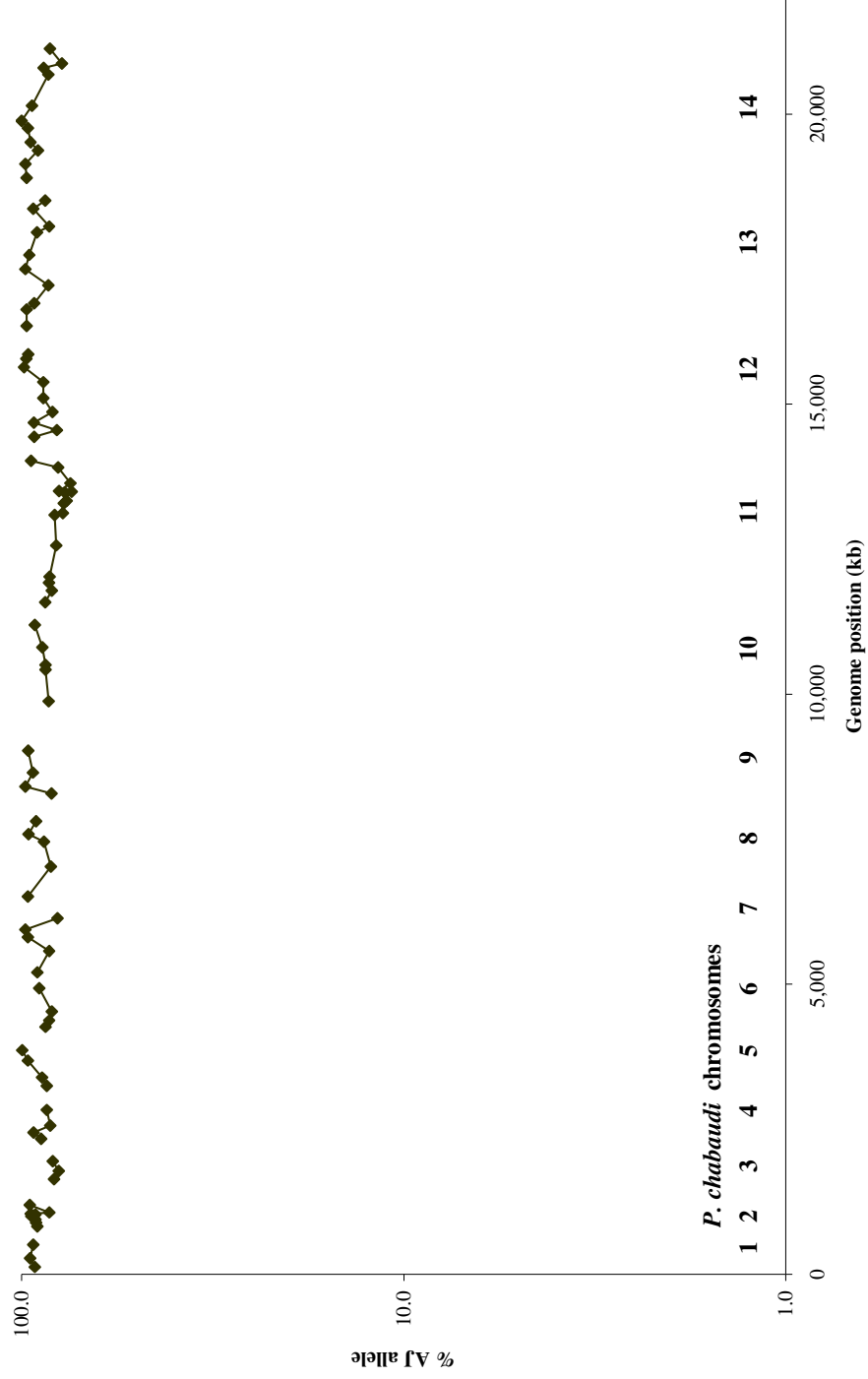


Figure 4.11 Selection 2 of AS-30CQ x AJ backcross – treated with 1.5 mg of CQ kg⁻¹ day⁻¹

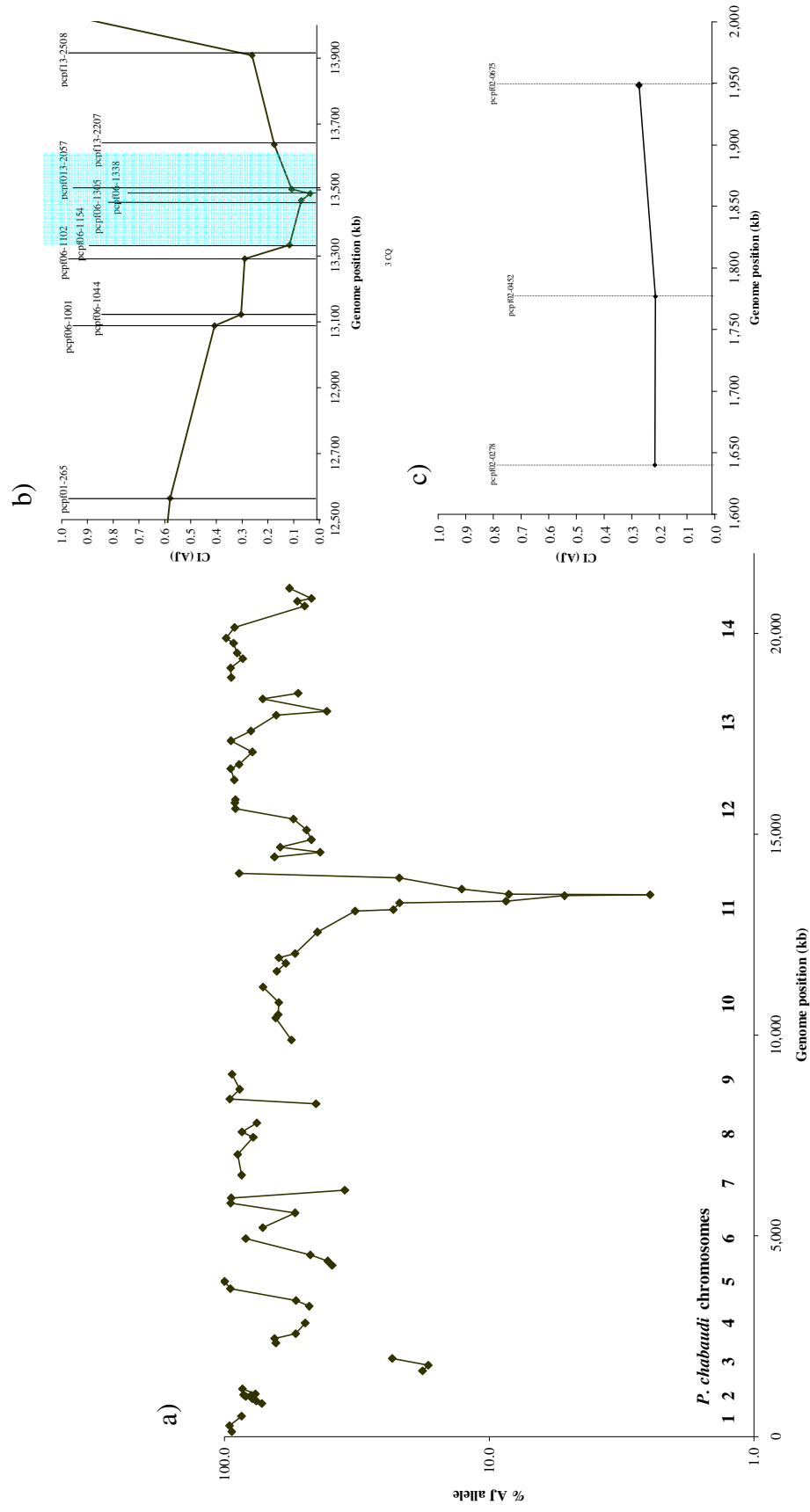


Figure 4.12 Selection 2 of AS-30CQ x AJ backcross – treated with 3 mg of CQ kg⁻¹ day⁻¹

a) the whole genome scan in log .scale, b) the region identified previously associated with low CQ resistance in AS-3CQ x AJ cross (Hunt et al., 2005) is shaded c) the selection valley on chromosome 3, Note that the whole genome data is presented in log scale while the selection valleys are not.

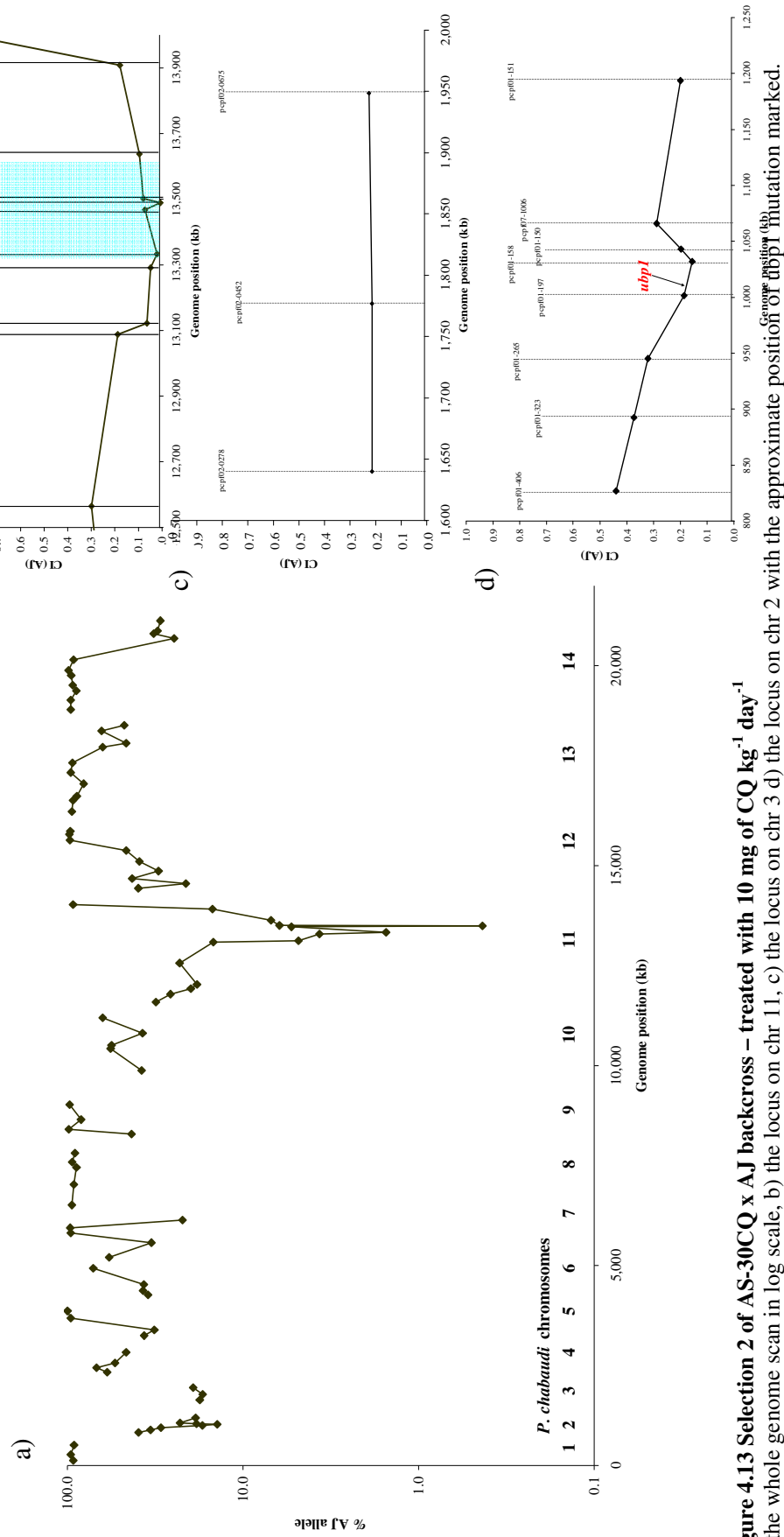


Figure 4.13 Selection 2 of AS-30CQ x AJ backcross – treated with 10 mg of CQ kg⁻¹ day⁻¹

a) the whole genome scan in log scale, b) the locus on chr 11, c) the locus on chr 3 d) the locus on chr 2 with the approximate position of *ubp1* mutation marked.

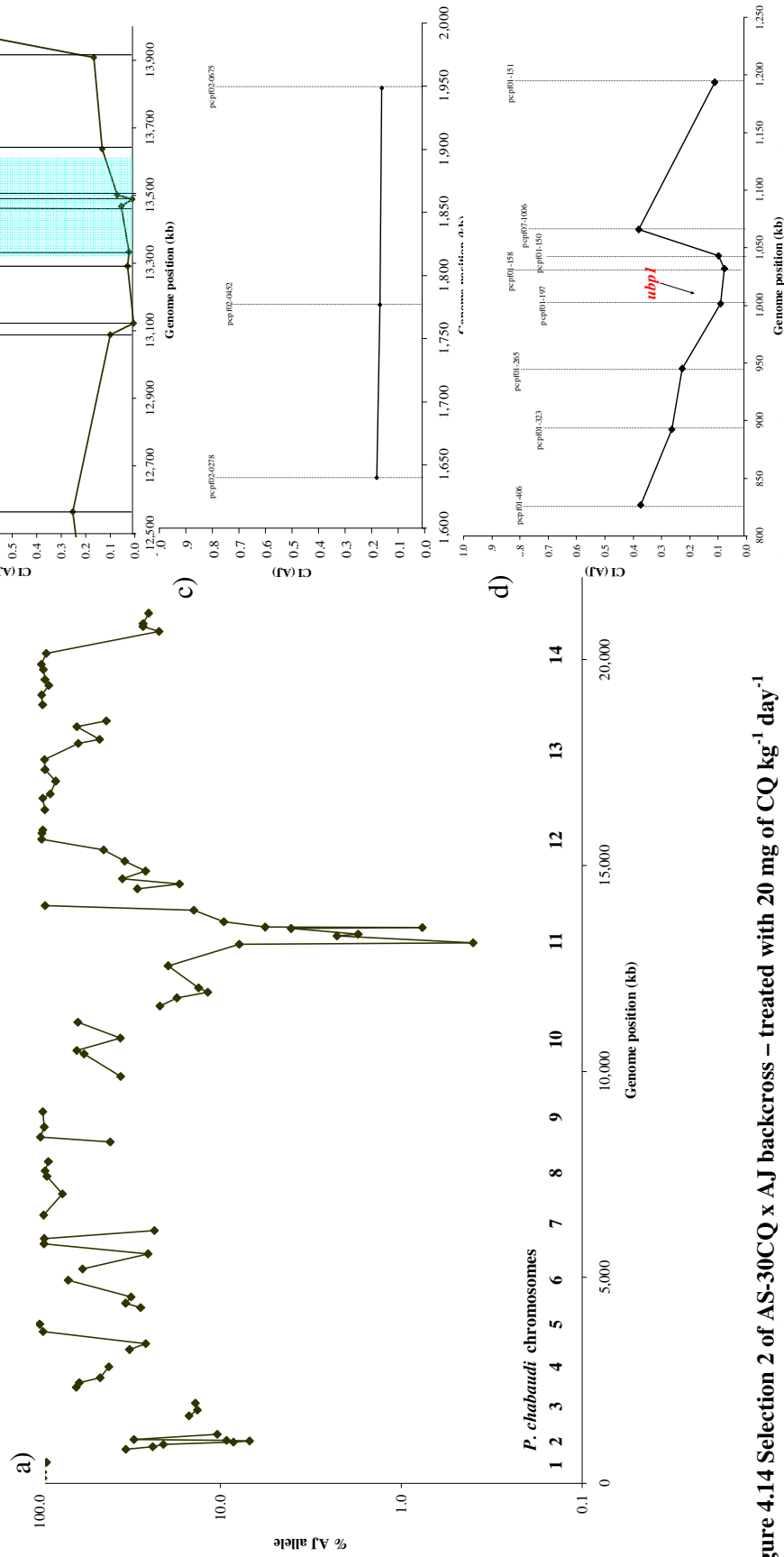


Figure 4.14 Selection 2 of AS-30CQ x AJ backcross – treated with 20 mg of CQ kg⁻¹ day⁻¹
a) the whole genome scan in log scale, b) the locus on chr 11, c) the locus on chr 2 with the approximate position of *ubp1* mutation marked.

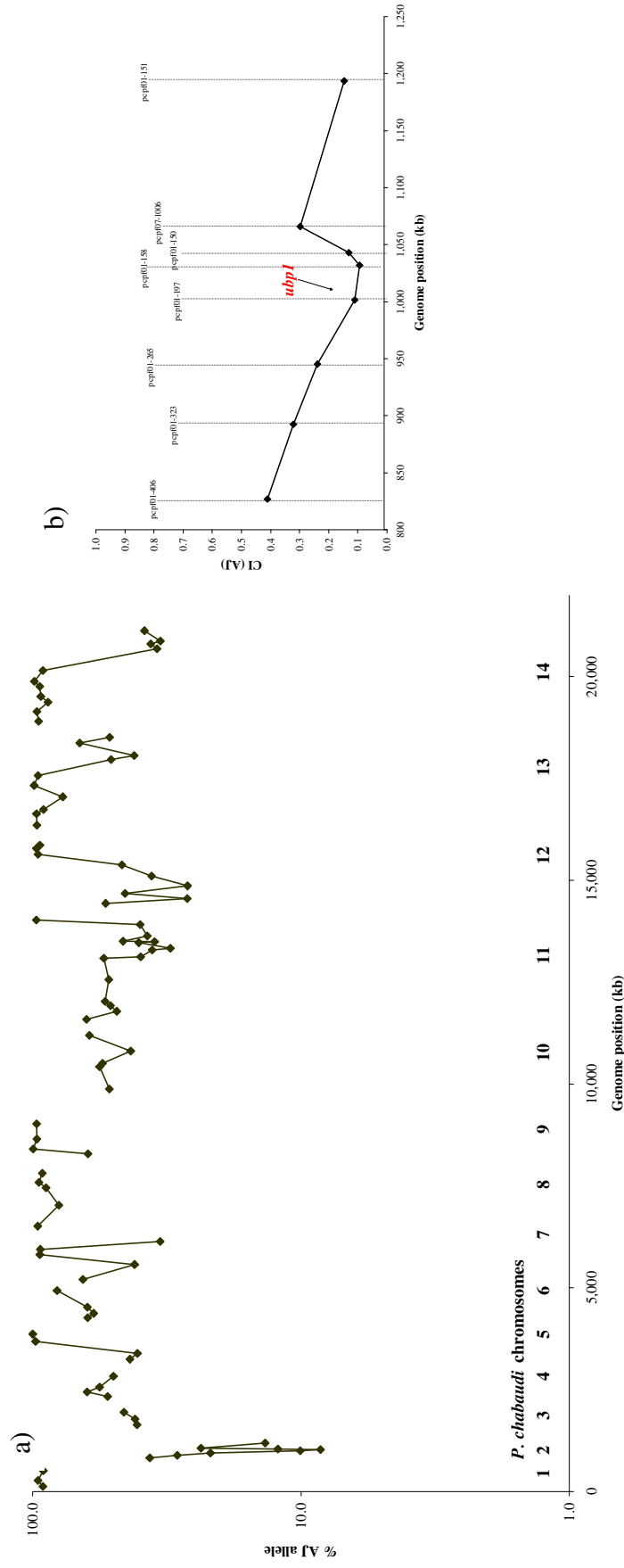


Figure 4.15 Selection 2 of AS-30CQ x AJ backcross – treated with 100 mg of ART kg⁻¹ day⁻¹
a) the whole genome scan in log scale, b) the locus on chr 2 with the approximate position of *ubp1* mutation marked.

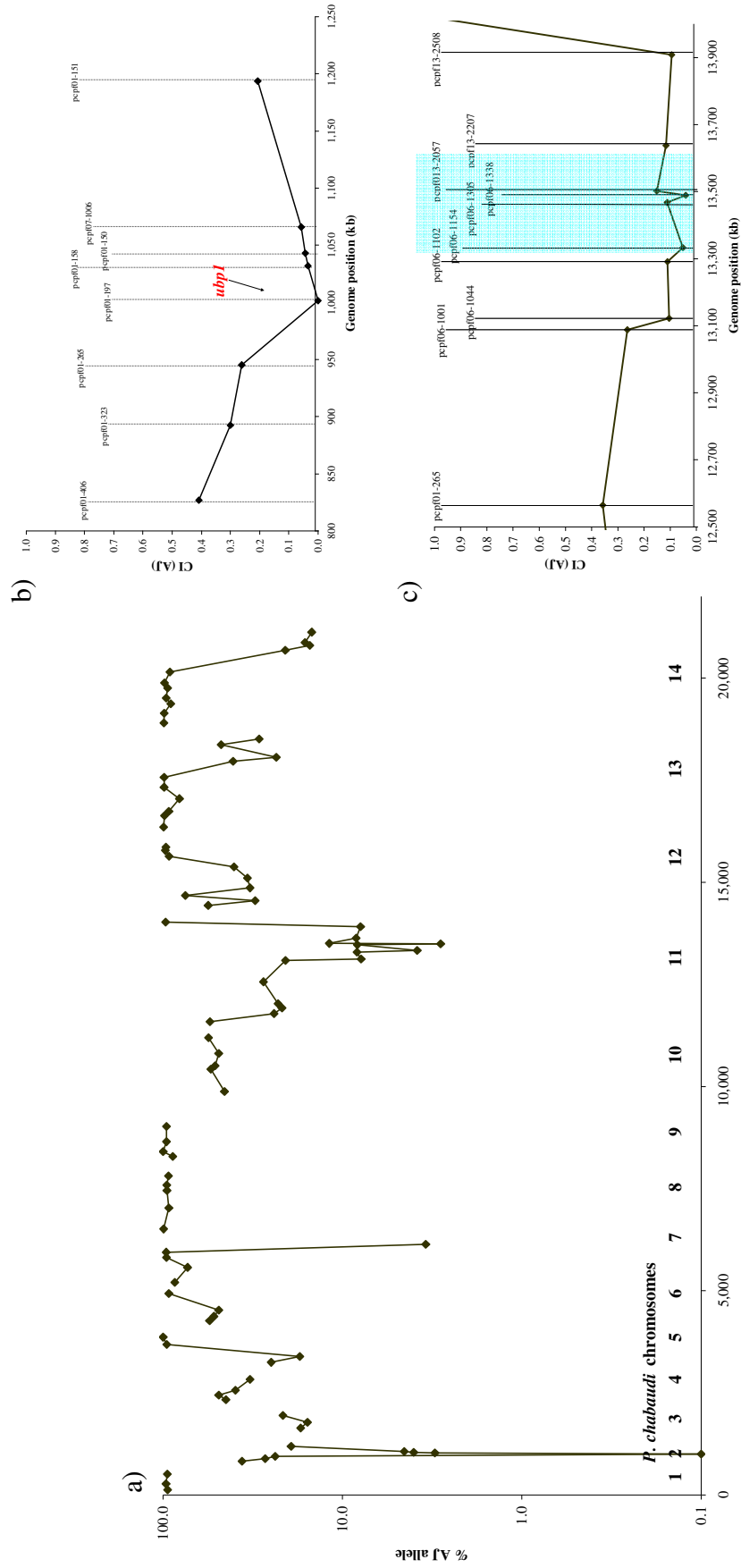


Figure 4.16 Selection 2 of AS-30CQ x AJ backcross – treated with 200 mg of ART kg⁻¹ day⁻¹

a) the whole genome scan in log scale, b) the locus on chr 11 with the potential region implicated in low CQ resistance marked, c) the locus on chr 2 with the approximate position of *ubp1* mutation marked, c) the locus on chr 11 with the potential region

4.5. Discussion

4.5.1. The improvement in the quality of data

In the discussion of the previous chapter (section 3.5.4 of Chapter 3), I expressed concerns regarding the quality of the data generated using the original AS-30CQ x AJ cross and the number of resistant recombinants in the final mixture. Therefore modifications of the original LGS protocol, presented in the introduction to this chapter, were introduced to generate and maintain the higher numbers of recombinants in the mixtures and to improve the selection conditions.

The initial results obtained during the backcross generation were encouraging. Firstly the study of the pre-selected progeny before the backcross revealed a well-balanced population containing both AJ and AS alleles as well as evidence of a reasonable number of recombinants. They also confirmed ideas regarding the extensive loss of resistant parasites in untreated infections. In fact, by following the original backcross protocol (the amplification of progeny without drug pressure and backcross to sensitive parent) the populations may have contained on average ~90% of AJ markers at most loci, generating very few recombinant parasites which are drug-resistant.

Secondly the numbers of oocysts generated during the backcross were around 8-fold greater than in the cross, therefore potentially supplying many more recombinants. Whether this was due to EPO treatment, to improved savoir-faire or to some other factors cannot be established.

As expected from these data, the significant improvement of data quality was observed during the analysis of the initial selection of the backcross (selection 1). Thus, the distributions of allele frequencies were less bimodal (allele proportions were well distributed over the range ~0% AJ – ~100% AJ), there was more evidence of linkage between adjacent markers (similarities in allele proportion) and there were fewer highly distinct regions of selection. In fact these data allowed me to dismiss a large number of the potential loci suggested in the original cross. It also mapped chromosome 11 as the main locus of resistance and suggested the involvement of chromosomes 3 and 2.

However further improvement was required. The variation between the neighbouring markers was still relatively high and this “background noise” did not allow resolution of the selection valleys *within* the chromosomes identified. Moreover all of the mapped selection valleys were relatively shallow: none of the markers within them were reduced below 20%. In the case of 1.5 mg of CQ kg⁻¹ day⁻¹ this result can be easily explained - the experiment described in chapter 2 (see Fig. 2.9) demonstrated that this concentration offers incomplete selection and allows some of the sensitive parasites to survive. Therefore we would expect only the moderate decrease in the frequency of AJ markers at the loci of resistance. However at higher concentrations the sensitive parasites were expected to have been eliminated and the selection should result in complete elimination of AJ markers at that locus.

There are a few possible explanations for this phenomenon. For example, inefficient drug treatment of one or more animals in the treatment group would reduce the apparent selection in the pooled material by allowing the growth. Alternatively, the proportion of the AJ alleles after the first (untreated) passage after the cross was already very high and the growth in the treated animals during selection was perhaps not long enough to change the proportion completely in favour of resistant allele. Yet another possibility was that (possibly due to the number of resistant parasites) the length of the infection was sufficient to allow the recrudescence of the very low numbers of potentially surviving sensitive parasites. Such recrudescence was observed even with very high CQ treatments in chapter 2 (see Fig 2.6).

Trying to further improve the quality of the data I therefore pre-selected the backcross progeny in order to increase the numbers of resistant parasites and subjected it to another series of selections (selection 2).

These resulted in a spectacular improvement in data quality resulting in a consistent pattern of selection valleys associated with CQ and ART resistance.

4.5.2. Three levels of resistance to chloroquine conferred by three loci?

Three loci of interest have been discovered during the treatment of the backcross with chloroquine –chr 11, chr 3 and chr 2.

A deep selection valley on chromosome 11 between markers pcpf 06-1154 and pcpf 13-2057 was the most prominent common feature of all the genome scans with the intermediate and high CQ treatments. It was also the *only* locus selected with the lowest used concentration ($1.5 \text{ mg kg}^{-1} \text{ day}^{-1}$) in one of the selections. Moreover during selection 2 some AJ alleles within this locus were *completely* or nearly completely selected in populations treated with intermediate and high CQ doses, indicating that it is necessary for the highest level of resistance. Interestingly, the same region had been previously identified as conferring CQ resistance during classical linkage analysis (in independent recombinant progeny clones) of the AS-3CQ x AJ cross (Carlton et al., 1998; Hunt et al., 2004a). Taken together, these data strongly suggest that chr 11 contains a mutation that appeared during the selection of AS-3CQ from AS-PYR1, that confers the low level of CQ resistance and seems to be necessary (but not sufficient) for the higher CQ resistance levels.

A secondary locus selected (although not completely) with intermediate and high level of CQ resistance was selected on chromosome 3. A mutation on chr 3 is therefore expected to confer higher levels of chloroquine resistance, possibly acting epistatically in addition to the main locus on chromosome 11. Therefore a mutation here is likely to appear later in the lineage, possibly between the low resistant AS-3CQ and intermediate-resistant AS-15CQ.

Finally chromosome 2 was selected in both selection 1 and 2, only with the highest level of CQ. The lack on selection on this locus with the intermediate concentration suggests that the mutation within it appeared during the selection for highest level of resistance (possibly between AS-15 CQ and AS-30CQ.) Interestingly, one mutation that appeared within that locus under chloroquine pressure (between AS-3CQ and AS-30CQ) has already been identified - a V2728Fsubstitution in the *ubp1* gene (Hunt et al., 2007). However the data presented here does not allow us to confirm or reject the theory that said mutation

contributes to the chloroquine resistance phenotype. Other genetic changes may be present on this chromosome and the *ubp1* mutation could be selected together with the critical mutation due to close genetic linkage.

Taken together, the LGS data from all the CQ selection experiments allowed the construction of a possible model of CQ resistance acquisition, in which only one locus would be responsible for each increase of the resistance (Fig. 4.17). Thus one locus, situated on chromosome 11, containing a mutation that appeared first in the lineage between AS-PYR1 and AS-3CQ, would be responsible for lowest level of resistance. Two loci (chromosomes 11 and 3) would contain mutations conferring an intermediate level of resistance to in AS-15CQ. Finally the third locus (on chromosome 2) would appear in AS-30CQ allowing it to achieve the highest level of resistance. This last mutation may be *ubp1* V2728F.

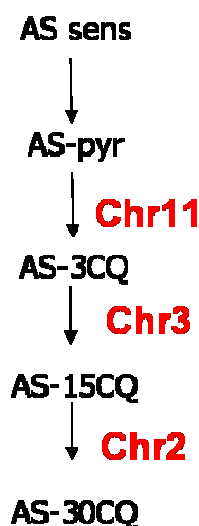


Figure 4.17 A model of how mutations may confer CQ resistance in AS lineage.

The strains are presented in black. The arrows are representing the selection and on red there are loci potentially mutated during the said selection.

It has to be noted however that this model is one of many possibilities. For example the mutation within three loci could have all appeared during the initial selection for AS-3CQ and the later increase of resistance would be caused either by additional mutations within the same loci/genes (similarly to acquisition of resistance to pyrimethamine (Lozovsky et al., 2009)) or many small effect mutations that would be undistinguishable from background noise. Alternatively the mutations within all

of the predicted loci could arise only in AS-30CQ and other genetic mechanisms are responsible for the lower level of resistance.

4.5.3. .Artemisinin resistance – confirmation of existence of the locus on chromosome 2

It has been previously reported that a locus on chromosome 2 may be responsible for artemisinin resistance (Hunt et al., 2007). However, at that time, no complete selection on this locus could be observed in the AS-ART x AJ cross. Also the appearance of a potential candidate mutation in the *ubp1* gene appeared not to correspond to the AS-30CQ phenotype (which was originally thought to be ART-sensitive).

During my work on CQ resistance however, re-phenotyping of some of the strains in the AS lineage (performed by Richard Fawcett and Louise Rodrigues) established that AS-30CQ also shows an artemisinin-resistance phenotype and that higher ART concentrations than those explored in the previous work should be used in selection experiments. These findings opened the opportunity to use the AS-30CQ x AJ cross for mapping of the loci of ART resistance in the lineage.

Now, using the AS-30CQ x AJ backcross and high artemisinin concentrations I have confirmed that in the resistant population almost all parasites carry resistant parental AS alleles on chromosome 2 thus establishing this locus as the main determinant of ART resistance in the lineage. The generated selection valley is very similar to that observed with the highest chloroquine concentration prompting the hypothesis that the same mutation could be responsible for both an increase of the resistance to chloroquine *and* appearance of resistance to artemisinin. This would explain why the artemisinin resistance seem to appear “out of nowhere” (without the prior exposure to artemisinin) in the AS lineage. The mutation in *ubp1* would be selected by CQ and then confer ART resistance in a multi-drug resistance phenotype. However, as explained previously, these genetic data do not allow us to confirm or reject the specific importance of *ubp1* mutations.

4.5.4. The possible locus of ART resistance on chr 11

Surprisingly with the highest artemisinin concentration, AS alleles on chromosome 11 were also selected, forming a selection valley similar to that obtained from CQ selected populations.

One possibility is that a locus on that chromosome (possibly identical with the main determinant of CQ resistance) is responsible for the higher level of artemisinin resistance. Therefore two very similar loci (on chr 2 and 11) would be selected by two different drugs (although in the reverse order). It would imply either strong structural/functional similarities between the two drugs (so far not supported by the literature) or the existence of “multidrug resistance” phenotype conferring the resistance to various drugs (similar to *pfmdr1* phenotype (Duraisingh and Cowman, 2005)).

There is, however another simple explanation of the data here. It is possible that the observed result is an artefact of the used experimental design. In particular case, the two selections are not totally independent – the parasites that are taken for artemisinin selection have previously been growing in untreated or chloroquine-treated mice. In this case, linkage disequilibrium can appear between the two loci as explained below.

Let us assume that in the absence of drug the mutation on chromosome 2 is not neutral and confers some cost to parasites such as slower growth etc. At the same time it increases the level of chloroquine resistance but only if found in the presence of mutation on chromosome 11. In that case, if parasites are grown in untreated animals those bearing the mutation on chromosome 2 will be at disadvantage. In chloroquine treated mice the parasites either with a mutation on chr 11 or those bearing mutations on both chr 2 and chr 11 will be selected. Those bearing mutations on chromosome 2 alone will carry the fitness cost and because they are not resistant to CQ, these parasites will be eliminated in both populations. So before selection, we combine the two populations in which the proportion of clones bearing the mutation on chromosome 2 alone is decreased, although many double mutants (chr 11/chr 2) are present. If we now treat these parasites with ART selecting for the locus on chromosome 2, we are likely to select many parasites bearing the mutation

on chr 11 although the latter is not directly involved in resistance. This would result in a genome scan similar to that observed in experiments presented here. At 100 mg of ART this effect wouldn't be as visible because the selection on chr 2 is not complete and therefore the putative selection valley on chr 11 would be very shallow and impossible to distinguish from the background noise (however it needs to be noted that with the exception of chr 2 the lowest data points on 4.15 graph belong to chr 11).

In summary this chapter has used genetic approaches to define the dominant genetic loci underlying resistance to CQ and ART. We expect to find mutations within these loci in the resistant parasite AS-30CQ relative to its sensitive precursors. The next chapter describes the search for such mutations.

5. The specific mutations connected with CQ and ART resistance

5.1. Introduction

The genetic experiments described in chapter 4 identified three loci involved in increasing levels of chloroquine resistance (on chromosomes 11, 3 and 2) and one locus involved in artemisinin resistance (on chromosome 2). All the identified loci however are relatively large. For instance the best resolved locus containing the main determinant of CQ resistance was mapped to ~300 kb fragment on chromosome 11. In the *P.chabaudi* genome, this translates to ~79 gene candidates (Table 5.1), many of which may possess functional or structural characteristics consistent with a role in CQ resistance, such as the involvement in intracellular transport, proteolysis, transcription control etc.

The narrowing of these loci using LGS would require additional cycles of backcrossing and selection, combined with an increase of marker density in the regions of interest. A significant amount of work would be involved and the improvement with each backcross would be smaller. Therefore a different approach was needed to map the candidate mutation within identified loci.

During the course of the project next-generation sequencing methodologies became available. We adopted the Illumina-Solexa platform - a novel, high-throughput, next generation sequencing system. This technology, based on the generation of short (36 or 50bp) reads allows the re-sequencing of the entire genome at high coverage in relatively short periods of time (less than 2 weeks for a *Plasmodium* size genome). Therefore two clones of different levels of resistance could be sequenced, mutations identified and all differences between them mapped and confirmed.

This chapter describes Solexa sequencing of the genomes of two clones from the AS lineage, namely AS-sens (the sensitive clone) and AS-30CQ (chloroquine and artemisinin resistant clone used as resistant parent in the genetic cross). These data are analysed and the critical genes connected with CQ and ART evaluated.

Gene_ID	Genomic position	Gene description
PCHAS_112600	chab11: 908,374 - 911,088 (-)	acetyl-CoA synthetase, putative
PCHAS_112610	chab11: 912,769 - 913,906 (+)	conserved Plasmodium protein, unknown function
PCHAS_112620	chab11: 914,737 - 915,255 (+)	6-pyruvoyltetrahydropterin synthase, putative
PCHAS_112630	chab11: 915,620 - 941,971 (-)	HECT-domain (ubiquitin-transferase), putative
PCHAS_112640	chab11: 946,368 - 954,095 (+)	protein kinase, putative
PCHAS_112650	chab11: 955,242 - 957,695 (-)	ethanolaminephosphotransferase, putative
PCHAS_112660	chab11: 960,061 - 961,923 (+)	conserved Plasmodium protein, unknown function
PCHAS_112670	chab11: 962,225 - 963,133 (-)	conserved Plasmodium protein, unknown function
PCHAS_112680	chab11: 963,922 - 965,632 (-)	conserved Plasmodium protein, unknown function
PCHAS_112690	chab11: 966,530 - 968,203 (+)	conserved Plasmodium protein, unknown function
PCHAS_112700	chab11: 968,714 - 970,825 (-)	glutamyl-tRNA(Gln) amidotransferase subunit B, putative
PCHAS_112710	chab11: 972,353 - 973,231 (-)	conserved Plasmodium protein, unknown function
PCHAS_112720	chab11: 973,861 - 975,876 (-)	conserved Plasmodium protein, unknown function
PCHAS_112730	chab11: 977,360 - 978,499 (-)	DNAJ protein, putative
PCHAS_112740	chab11: 980,219 - 984,359 (+)	RAP protein, putative
PCHAS_112750	chab11: 984,729 - 985,157 (-)	conserved Plasmodium protein, unknown function
PCHAS_112760	chab11: 986,826 - 989,045 (+)	phosphatidylcholine-sterol acyltransferase precursor, putative
PCHAS_112770	chab11: 991,735 - 993,267 (+)	RNA binding protein, putative
PCHAS_112780	chab11: 994,493 - 996,430 (-)	amino acid transporter, putative
PCHAS_112790	chab11: 997,701 - 998,547 (+)	conserved Plasmodium protein, unknown function
PCHAS_112800	chab11: 1,000,593 - 1,018,990 (+)	SET-domain protein, putative
PCHAS_112810	chab11: 1,019,613 - 1,022,894 (-)	cullin-like protein, putative
PCHAS_112820	chab11: 1,024,883 - 1,026,069 (+)	phosphatidylinositol/phosphatidylcholine transfer protein, putative
PCHAS_112830	chab11: 1,026,414 - 1,028,270 (-)	CPW-WPC family protein, putative
PCHAS_112840	chab11: 1,029,541 - 1,031,358 (-)	conserved Plasmodium protein, unknown function
PCHAS_112850	chab11: 1,033,102 - 1,033,993 (+)	conserved Plasmodium protein, unknown function
PCHAS_112860	chab11: 1,034,542 - 1,043,380 (-)	DNA polymerase epsilon, putative
PCHAS_112870	chab11: 1,045,450 - 1,046,631 (-)	conserved Plasmodium protein, unknown function
PCHAS_112880	chab11: 1,048,816 - 1,050,264 (+)	microtubule-associated protein ytm1 homologue, putative
PCHAS_112890	chab11: 1,051,795 - 1,054,489 (+)	conserved Plasmodium protein, unknown function
PCHAS_112900	chab11: 1,055,834 - 1,057,801 (+)	conserved Plasmodium protein, unknown function
PCHAS_112910	chab11: 1,058,902 - 1,061,225 (+)	conserved Plasmodium protein, unknown function
PCHAS_112920	chab11: 1,061,530 - 1,063,447 (-)	DEAD/DEAH box ATP-dependent RNA helicase, putative
PCHAS_112930	chab11: 1,065,707 - 1,066,549 (+)	conserved Plasmodium protein, unknown function
PCHAS_112940	chab11: 1,067,224 - 1,072,840 (-)	conserved Plasmodium protein, unknown function
PCHAS_112950	chab11: 1,073,888 - 1,076,767 (+)	ran-binding protein, putative
PCHAS_112960	chab11: 1,077,741 - 1,078,724 (+)	Transcription factor Tfb4, putative
PCHAS_112970	chab11: 1,079,719 - 1,080,683 (-)	ER lumen protein retaining receptor,

Gene_ID	Genomic position	Gene description
PCHAS_112980	chab11: 1,083,151 - 1,083,657 (+)	conserved Plasmodium protein, unknown function
PCHAS_112990	chab11: 1,084,283 - 1,085,118 (-)	proteasome subunit, putative
PCHAS_113000	chab11: 1,086,560 - 1,087,600 (+)	proteasome subunit, putative
PCHAS_113010	chab11: 1,088,162 - 1,089,721 (-)	conserved Plasmodium protein, unknown function
PCHAS_113020	chab11: 1,090,509 - 1,091,522 (-)	conserved Plasmodium protein, unknown function
PCHAS_113030	chab11: 1,092,442 - 1,094,799 (-)	inositol-polyphosphate 5-phosphatase, putative
PCHAS_113040	chab11: 1,095,860 - 1,098,986 (-)	methyltransferase, putative
PCHAS_113050	chab11: 1,100,447 - 1,101,532 (-)	V-type ATPase, putative
PCHAS_113060	chab11: 1,104,058 - 1,105,503 (+)	adenylosuccinate synthetase, putative
PCHAS_113070	chab11: 1,106,044 - 1,107,116 (-)	60S ribosomal protein L7-2, putative
PCHAS_113080	chab11: 1,108,004 - 1,109,113 (-)	conserved Plasmodium protein, unknown function
PCHAS_113090	chab11: 1,109,761 - 1,111,560 (-)	protease, caspase family, putative
PCHAS_113100	chab11: 1,112,862 - 1,114,550 (+)	conserved Plasmodium protein, unknown function
PCHAS_113110	chab11: 1,116,520 - 1,119,342 (+)	replication licensing factor, putative
PCHAS_113120	chab11: 1,119,557 - 1,122,079 (-)	RAP protein, putative
PCHAS_113130	chab11: 1,123,017 - 1,124,552 (+)	histone-lysine N-methyltransferase, putative
PCHAS_113132	chab11: 1,125,573 - 1,125,645 (-)	transfer RNA-Thr
PCHAS_113150	chab11: 1,129,758 - 1,133,054 (-)	phosphatase, putative
PCHAS_113160	chab11: 1,138,618 - 1,139,301 (+)	splicing factor 3b subunit, putative
PCHAS_113170	chab11: 1,140,112 - 1,141,305 (+)	conserved Plasmodium protein, unknown function
PCHAS_113180	chab11: 1,142,420 - 1,146,187 (+)	conserved Plasmodium protein, unknown function
PCHAS_113190	chab11: 1,147,086 - 1,149,743 (-)	conserved Plasmodium protein, unknown function
PCHAS_113200	chab11: 1,151,156 - 1,151,879 (-)	mitochondrial inner membrane translocase, putative
PCHAS_113210	chab11: 1,154,436 - 1,155,927 (+)	ubiquitin conjugating enzyme, putative
PCHAS_113220	chab11: 1,157,481 - 1,160,420 (+)	phosphatase 2A regulatory subunit-related protein, putative
PCHAS_113230	chab11: 1,161,499 - 1,162,197 (+)	conserved Plasmodium protein, unknown function
PCHAS_113240	chab11: 1,162,753 - 1,164,192 (-)	regulator of chromosome condensation, putative
PCHAS_113250	chab11: 1,166,278 - 1,167,579 (+)	DNAJ protein, putative
PCHAS_113260	chab11: 1,169,869 - 1,180,434 (+)	serine/threonine protein kinase, putative
PCHAS_113270	chab11: 1,181,572 - 1,183,096 (-)	Protein Kinase 5, putative
PCHAS_113280	chab11: 1,183,750 - 1,185,081 (-)	elongation factor 1 alpha, putative
PCHAS_113290	chab11: 1,186,171 - 1,187,502 (+)	elongation factor 1 alpha, putative
PCHAS_113300	chab11: 1,188,670 - 1,191,204 (+)	glutamate - tRNA ligase, putative
PCHAS_113310	chab11: 1,192,128 - 1,192,493 (+)	dynein light chain, putative
PCHAS_113320	chab11: 1,193,606 - 1,196,710 (+)	conserved Plasmodium protein, unknown function
PCHAS_113330	chab11: 1,198,016 - 1,200,643 (+)	DNA helicase, putative
PCHAS_113340	chab11: 1,201,367 - 1,202,126 (+)	conserved Plasmodium protein, unknown function
PCHAS_113350	chab11: 1,202,795 - 1,206,907 (-)	conserved Plasmodium protein, unknown function
PCHAS_113360	chab11: 1,208,050 - 1,210,123 (+)	TCP-1/cpn60 chaperonin, putative
PCHAS_113370	chab11: 1,211,265 - 1,212,774 (+)	pyrroline carboxylate reductase, putative

Table 5.1 The initial list of gene candidates on chromosome 11

This list includes all the genes present within the predicted selection valley according to PlasmoDB database (<http://plasmodb.org>).

5.2. The experimental design

DNA from clones AS-sens and AS-30CQ was sequenced at high coverage (around 20x for AS-sens and >30x for AS-30CQ) in short DNA sequence strings. All the reads were mapped onto a pre-existing fully assembled and annotated *P. chabaudi* AS reference genome sequence (AS-WTSI, ftp://ftp.sanger.ac.uk/pub/pathogens/P_chabaudi/September_2009 assembly) allowing the reassembly of the genome sequences of both AS-sens and AS-30CQ. This analysis was performed in parallel, using MAQ (<http://maq.sourceforge.net>) (Li et al., 2008) and SSAHA2 (<http://www.sanger.ac.uk/resources/software/ssaha2/>) (Ning et al., 2001) software packages.

The data were then analysed for SNPs and indels and the list of all genetic changes between the two strains compiled. This list was then filtered to remove poor quality calls on the basis of their associated quality indices and properties (e.g. read coverage, Phred-like scores, uniqueness and mapping quality, see section 5.3.3. for detailed description). Then the comparison of the data from AS-sens and AS-30CQ sequences allowed the identification of changes that appeared within AS lineage, during the selection for drug resistance. Once the putative mutations were identified, dideoxy sequencing was used to validate the results of bioinformatics analysis.

Later all confirmed changes were sequenced in all the other strains in the AS lineage in order to establish the place of their appearance within the lineage.

In order to be considered as a potential cause of drug resistance, the mutations had to fit two major criteria:

- they are situated within the limits of the predicted locus of resistance (identified in chapter 4 during genetic analysis)
- they appear during the chloroquine selection along with the change in chloroquine resistance phenotype (they could not be present in AS-sens strain)

Critical mutations affecting drug responses were expected to be non-synonymous as they are most likely to alter the function of the expressed protein in the parasite. However, it was important to note that synonymous and intergenic point mutations, indels and copy number variations (CNVs) could also mediate drug resistance by changing the levels of expression of a given protein and must therefore be considered as potential resistance determinants.

Once putative resistance conferring mutations were identified, proportional sequencing was used to assess the level of these mutations in the selected backcross progeny. This confirms that they were amongst the lowest points in the selection valleys and therefore the likely cause of the observed selection. Additionally the allele proportions in populations treated with different drug concentration could be used to estimate the strength of selections.

Finally, the properties of the candidate genes and proteins including their function, localisation, secondary structure, conservation in the *Plasmodium* genus and their possible role in the resistance mechanisms were investigated.

5.3. Materials and Methods

5.3.1. The preparation of DNA

The parasites used for this experiment were described in section 1.7.2. of Chapter 1. Both parasites were grown in groups of 20 male CBA mice, infected and monitored as described in section 2.3 of Chapter 2. DNA isolation was performed as described in section 3.3.7. of Chapter 3. DNA was quantitated spectrophotometrically and its quality and quantity confirmed by agarose gel electrophoresis.

5.3.2. Illumina-Solexa genome re-sequencing:

Both DNA samples were sequenced using Illumina-Solexa genome analyser (http://www.illumina.com/systems/genome_analyzer.ilmn) by Gene Pool sequencing service of the University of Edinburgh (<http://genepool.bio.ed.ac.uk/>). AS-sens and AS-30CQ were sequenced using 4 and 6 lanes of a flowcell respectively. 36 bp (AS-sens) or 50 bp (AS-30CQ) single-ended reads were generated.

5.3.3. The bioinformatic analysis of Solexa genome re-sequencing data

This part of the work was performed by Dr Axel Martinelli, members of the GenePool sequencing service (Marian Thomson, Urmi Trivedi, Sujai Kumar and Prof. Mark Blaxter) and Dr Paul Hunt.

5.3.3.1. Genome reassembly and SNP identification using MAQ software

In MAQ the obtained reads were mapped against the AS reference sequence ([AS-WTSI, ftp://ftp.sanger.ac.uk/pub/pathogens/P_chabaudi/September 2009_assembly](http://ftp.sanger.ac.uk/pub/pathogens/P_chabaudi/September_2009_assembly)) using the “easyrun” command, run with default parameters. This automatically generated a list of SNPs between a clone and the reference sequence (as a “cns.final.snp” file) filtered using following criteria:

- minimum coverage of 3 (the base needed to be covered by at least 3 reads to be called as SNP)

- minimal Phred-like score of 30 (the Phred-like score is calculated on the basis of Phred-scores assigned by Solexa sequencing software to each sequenced base and reflecting the certainty of the call)

This list of SNPs was then further filtered, using three custom perl scripts. The first two were applied to both AS-sens and AS-30CQ dataset and used to remove the non-unique and heterozygous calls. Non-unique calls are the SNPs identified by the reads aligned to more than one place to the genome and therefore likely to be incorrectly positioned. Heterozygous calls are identified when two possible bases were identified on the given position with the “recessive” base constituting at least 20% of the calls. The major base was always identical to the one in reference sequence. The minor base represented either incorrectly mapped reads, or a polymorphism present *within* the sequenced strain. The third script was used to filter the AS30CQ calls against the AS-sens one generating the list of SNPs that appeared during the selection of drug resistant lineage. All of the identified SNPs were visualised using “mapview” command and further inspection.

Locations spanning SNPs were subsequently characterized using BLAST searches against the PlasmoDB (<http://plasmodb.org>) and GeneDB (<http://www.genedb.org>) databases.

5.3.3.2. Genome reassembly and SNP/indels identification using SSAHA2 software

In SSAHA2 the reads from both strains were aligned against AS-WTSI using the “pileup” command for single end reads with the appropriate options as described in ftp://ftp.sanger.ac.uk/pub/zn1/ssaha_pileup/ssaha_pileup-readme. It generated three files calling snp (*.snp), 1-3bp deletions (*.del) and 1-3bp insertions (*.ins). They were all combined in Artemis-compatible “*.gff” file and divided according to their chromosomal location (a separate file was generated for each chromosome). The data from these files were filtered according to the following criteria: for the SNPs the minimal coverage of >3 was required and heterozygous calls were removed. For indels only the calls called by over 50% of the covering reads were kept. The resulting list of mutations was filtered against AS-sens.

The “gff” files were then imported into Artemis (<http://www.sanger.ac.uk/resources/software/artemis/>) (Carver et al., 2008), a DNA sequence viewer which allowed the visual inspection, identification and evaluation of these mutations on the fully annotated genome. Further information was obtained by consulting the PlasmoDB and GeneDB.

5.3.3.3.Mapping the indels and CNV (copy number variations) between the two strains

As both MAQ and SSAHA2 don't have an automatic algorithm for identification of larger deletions (>3bp) and CNV, a separate approach, based on the comparison of local read coverage in AS-30CQ and AS-sens was adopted using both MAQ and SSAHA2 outputs and custom designed script. This method is based on the assumption that insertions and deletions result in a reduction of read coverage that can be measured, and also upon the observation that (relative) read-coverage varies across the genome but in a regular way (in different clones) (Axel Martinelli, personal communication).

The datasets generated by both MAQ and SSAHA2 were used for that analysis. For MAQ data this required running the “pileup” command on both the AS-30CQ and AS-sens reads using the variable “-q 1” (which excludes reads with a mapping quality <1). The pileup files thus produced were then used for coverage analysis. For SSAHA2 the output of the “pileup” command was used directly for coverage analysis. The relative coverage at any base in one clone was defined as the actual coverage relative to the overall (genome-wide) coverage. The comparative coverage was defined as the ratio of the relative coverage in the mutant clone relative to that of AS-sens. Differences in coverage were evaluated separately for potential CNVs and potential large indels using custom made scripts. A potential CNV was defined as an increase in comparative coverage of at least 1.5-fold over a stretch of at least 200 contiguous bp. A potential large indel was defined as a reduction in comparative coverage of at least 75% in the mutant clone compared to AS-sens over at least 10 contiguous bp. Additionally, a minimum coverage of at least 20-fold in clone AS-sens was required, in order to remove an excess of false positive calls due to areas with insufficient coverage.

All suspected indels were visually confirmed using MAQ and Artemis software.

5.3.3.4. Calculation of numbers of mapped reads and genome coverage

The number and proportion of reads mapped in MAQ was calculated automatically. For the calculation of genome coverage, the output of the “pileup” command used to calculate the overall coverage, using a custom designed script.

For SSAHA2 a custom-made perl script was used to calculate the number of mapped reads. For coverage analysis, an “awk” script (supplied by Thomas D. Otto, Sanger Institute) was used to produce a plot from the “pileup” file. This plot was read by the ARTEMIS software which automatically produced the average coverage for each chromosome. The average coverage of the whole genome was based on the average of the coverage for all chromosomes, adjusted by the size of each individual chromosome.

5.3.4. Dideoxy sequencing

To confirm the bioinformatics predictions, a set of PCR primers was designed.. The standard PCR protocol was used to amplify fragments containing mutations of interest (Appendix B). The list of the primers used for mutations investigated in this chapter is presented in Table 5.2.

5.3.5. Proportional sequencing

The primers designed for the sequencing of the mutations (Table 5.2) were also used to confirm the proportions of parental alleles by proportional sequencing. Amplified products from the selection 2 of the backcross (Fig 4.9 in chapter 4) were sequenced and analysed as described in section 2.3.5. of chapter 4.

Chr.	Position	Type of change	Nucleotide change or indel size	Left primer	Right primer
2	216,954	SNP	C/A	CTTTTGTTTCTCATAATCCAG	GGTTTAGGTTTGTGCAAGAA
3	70,553	SNP	G/T	CTGGAACAAGCACAAACAGC	TTCATTGGTGCTCCTGTATC
3	180,102	SNP	G/A	GAAACTAACAGGCCCCACAC	TAGCACAATGGGAAAAAATTAG
3	180,086	SNP	A/C		
3	474,123	SNP	C/A	GAAACTAACAGGCCCCACAC	TAGCACAATGGGAAAAAATTAG
7	876,901	Deletion	34bp	CGTATAAAGGCTGTGACAAC	CAGGTTCGTTTGTACATACC
7	876,917	SNP	A/G		
7	876,919	SNP	C/A		
7	994,546	SNP	G/A	TTTTAAGGGACTTGGGAATG	ATAAATCTTCATTGATATCTGG
10	290,661	Deletion	1bp	CGTGTGTGCTTTATCTTGTTTA	CATTATTTTGTTCGGTTTC
10	634,932	SNP	T/C	GAGACGATGAGTGTTACTTTCC	AATAGTACCTGTGTGTGTTAGTCGC
11	760,507	Deletion	1bp	ACTCAAACCAAGAAAAACAAAAG	CGCTACCACTACGAATACT
11	432,447	SNP	T/C	AAACGTCCTAAAAAGTGAAGG	AACTTTGATGTGAAATGACATAGC
11	946,436	SNP	G/T	CACCATATAGCAATCATTAACG	CATTTTCATCATTTGTGTGTAC
11	946,392	SNP	C/T		
11	996,332	SNP	G/T	GGAGTGATAAGAATGATAATATGAGG	GTGTTCAAAGTGCTGACATGC
11	1,013,413	SNP	C/G	AAATAAAACAGGCCCATACAATC	AGAAATAGTGCAGCAAACTAAGC
13	113,002	SNP	G/A	TGCAGATTTTGTGTTGTTTCAG	TTGTAGATAGTGCCTCATTCA
14	936,945	SNP	T/G	CTTCACTATAAAATATGCTGTTC	CATGCCGATACTTATTATATGTC

Table 5.2 Primers used for validation of sequencing mutations.

The primers marked with * were also used for the proportional sequencing.

5.3.6. Protein characterisation

The publicly available PlasmoDB database (<http://plasmodb.org/>) was used to obtain the basic information concerning the predicted gene structure, coding sequence, possible function and expression profile. PlasmoDB was also used to find the sequences of the predicted proteins and their homologues in other *Plasmodium* spp. genomes. The Clustal W2 software (<http://www.ebi.ac.uk/Tools/clustalw2/index.html>) was used to generate alignments between orthologues. The analysis of secondary structure was performed using TMHMM (<http://www.cbs.dtu.dk/services/TMHMM/>) and Jpred 3 (<http://www.compbio.dundee.ac.uk/www-jpred/>).

5.4. Results

5.4.1. The genome re-sequencing and Solexa data analysis

5.4.1.1. Sequencing and initial data analysis:

All animals used for DNA amplification developed parasitaemia as expected and DNA was isolated on day 6. Both samples were sequenced as described in section 5.3.2. Given the size of the *Plasmodium* genome, such sequencing should result in the approximate coverage (number of times each base was sequenced) of 20X and 30X respectively.

The initial data analysis (Table 5.3) showed that with both the MAQ and SSAHA2 software the vast majority of reads could be mapped against the reference genome and the overall coverage for both AS-sens and AS-30CQ was better than expected (40x and 75x respectively). Moreover, in both strains the majority of the genome (~98%) was covered with reads with the coverage of at least 5X which indicated that only a very small portion of the genome was not covered by analysis. Additionally, the comparison of coverage on particular chromosomes in AS-30CQ (Table 5.4) indicates that the part of the genome that could not be covered with uniquely mapped reads was mostly confined to the contigs not assigned to any of the chromosomes (“bin”) and therefore likely to be a part of the low complexity regions of small significance (e.g. telomeric and sub-telomeric sequences).

Clone analysed	AS-sens	AS-30CQ
Amount of DNA used for sequencing	2 µg	2 µg
Number of lanes on a flowcell	4	6
Read length	36	41
Nominal coverage (calculated based on the minimal amount of data generated with each lane)	20X	30X
% reads allocated to sequence (MAQ)	90.83%	96.60%
% reads allocated to sequence (SSAHA)	93%	97%
Actual coverage (calculated by MAQ)	39.28X	77.91X
Actual coverage (calculated by SSAHA2)	39.48X	75.86X
% of the genome with coverage >10	95.86	96.96%

Table 5.3 The summary of Solexa sequencing metrics of AS-sens and AS-30CQ strain.

Chromosome	% coverage ≥10
Chr01	96.16%
Chr02	98.92%
Chr03	96.32%
Chr04	97.48%
Chr05	97.71%
Chr06	98.04%
Chr07	96.61%
Chr08	99.27%
Chr09	99.39%
Chr10	98.11%
Chr11	98.78%
Chr12	98.88%
Chr13	98.74%
Chr14	98.84%
Total (chrs only)	98.31%
bin	71.69%
Total	96.96%

Table 5.4 The summary of the coverage of AS-30CQ by chromosome in SSAHA2.
“bin” represents the contigs that were not mapped to any of the chromosomes

5.4.1.2. The identification of SNPs between the two strains

This analysis was performed in parallel using MAQ and SSAHA2 software.

Initial mapping of the reads against reference sequence, using default parameters of MAQ software (coverage >3 and Phred-like score >30), resulted in identification of hundreds of SNPs between each of the strains and reference sequence (see Table 5.5). The removal of non-unique and heterozygous calls however (as described in section 5.3.3.1.), reduced the SNP numbers to 34 and 41 in AS-sens and AS-30CQ respectively. The majority of those were shared between AS-sens and AS-30CQ and therefore either appeared very early in the history of AS-sens strain or were due to the errors in reference sequence and thus were unlikely to be connected with CQ/ART resistance. The remaining 17 SNPs that were detected in AS-30CQ strain only are presented in Table 5.6.

After the visual verification 7 out of those 17 were kept for further analysis. Two of the others (the ones chr 7) were discarded because they were situated on the edge of predicted deletion which caused the visible errors of read mapping and low coverage in this region. For the remaining 8, the analysis of AS-sens sequence revealed that they were present in this strain and were not called as SNPs between AS-sens and the reference sequence only because of poor quality and/or read coverage. The rejected SNPs were characterized by lower Phred-score than the confirmed ones suggesting that they were covered by lower quality, uncertain reads.

A similar analysis was performed using SSAHA2 software. First, the reads were uniquely mapped to the reference sequence without any filtering criteria which resulted in the generation of a large list of SNPs for each of the chromosomes (over 4000 for the whole genome, data not shown). The subsequent filtering for minimal coverage of 3, removal of heterozygous calls and comparison with AS-sens (as detailed in section 5.3.3.2.), reduced this list to 12 proposed SNPs between the two strains (Table 5.7). 7 of those were characterised by high coverage and quality score (value assigned by SSAHA2 and reflecting the probability that the SNP was called correctly) close to the maximal value of 99 percent. They were the same ones as

identified with MAQ and therefore judged to be likely mutations. The other 5 were all characterised by low coverage and very low quality score (<20) and lack of visual confirmation and therefore were judged as unlikely.

In summary 7 high quality SNPs were identified using both approaches. Additionally each of the two approaches produced few additional low-likelihood candidates with substantially lower quality that were considered to be false positives.

	AS-sens	AS-30CQ
Initially identified SNP	836	699
After removing heterozygous calls	54	59
After removing non-unique calls	34	41
Filter against AS-sens	-	17
After visual inspection	-	7

Table 5.5 Number of point mutations identified on each step of analysis with MAQ software

“Initially identified SNPs” refers to the MAQ software output. Heterozygous calls are the ones where two possible bases are called in particular position. Non-unique call refers to the mutation identified by the reads **mapping to more than** one place in the genome.

	Chr.	Position	Ref ^a	AS-sens		AS-30CQ		Phred-like score ^d	Reason for rejection
				Cove ^b rage	Base ^c	Cove ^b rage	Base ^c		
1	bin	107946	C	6	A	11	A	59	present in as-sens
2	bin	145080	C	8	T	54	T	146	present in as-sens
3	bin	208863	A	7	G	19	G	83	present in as-sens
4	chr02	216954	C	42	C	80	A	249	-
5	chr03	70553	G	28	G	123	T	255	-
6	chr03	474123	C	40	C	59	A	205	-
7	chr07	876917	A	16	A	16	A	72	proxy to indel
8	chr07	876919	C	23	C	23	C	91	proxy to indel
9	chr07	994546	G	66	G	92	A	255	-
10	chr10	634932	T	22	T	38	C	141	-
11	chr11	996332	G	59	G	117	T	255	-
12	chr13	1792583	C	1	T	9	T	54	present in as-sens
13	chr13	690908	A	10	T	22	T	93	present in as-sens
14	chr13	1063028	C	1	A	6	A	44	present in as-sens
15	chr13	1063060	C	3	A	4	A	39	present in as-sens
16	chr13	1063059	T	3	C	4	C	39	present in as-sens
17	chr14	936945	T	14	T	27	G	99	-

Table 5.6 The point mutations identified with MAQ software and the results of their visual confirmation.

The likely SNPs are marked in green, the unlikely ones are grey and the reason for rejection is given.

a) the base on the described position in reference sequence (AS-WISI) b) number of reads covering the given position in MAQ assembly, c) a base identified by visual analysis, d) a score assigned to each SNP identified by MAQ software in AS-30CQ, reflecting the certainty of the call (max.= 255)

	Chromosome	Position	Analysis ^a	AS-sens		AS-30CQ		SSAHA Quality Score ^c
				Base	Cove ^b rage	Base	Cove rage	
1	2	216,954	SSAHA/MAQ	C	42	A	72	99
2	3	70,553	SSAHA/MAQ	G	28	T	61	99
3	3	474,123	SSAHA/MAQ	C	40	A	57	99
4	5	681,914	SSAHA	T	3	A	4	6
5	7	994,546	SSAHA/MAQ	G	66	A	87	99
6	10	634,932	SSAHA/MAQ	T	22	C	37	99
7	11	996,332	SSAHA/MAQ	G	59	T	111	99
8	13	113,002	SSAHA	G	2	A	4	8
9	14	936,945	SSAHA/MAQ	T	14	G	26	92
10	bin	116,902	SSAHA	A	34	C	25	16
11	bin	167,611	SSAHA	T	1	C	4	15
12	bin	221,856	SSAHA	T	20	G	6	2

Table 5.7 The point mutations identified with SSAHA2 software.

The calls rejected because of the low read depth, low quality score and lack of visual confirmation are marked in grey. a) the software calling given SNP, b) number of reads calling given SNP in SSAHA2 assembly, c) quality score assigned by SSAHA software, representing the probability that given SNP is not a false positive

5.4.1.3. The identification of indels between the two strains

A separate approach was used to map small (1-3bp) indels and larger indels and CNV. The summary of the results is presented in Table 5.8.

For identification of small indels (<3bp) SSAHA2 internal algorithm was used based on identifying the 1 to 3 bp indels within the reads (see section 5.3.3.2.). It yielded a list of 30 potential indels present between AS-sens and AS-30CQ. However none of them could be confirmed by visual analysis and the majority was characterized by low coverage (<10) and/or was called only by some of the covering reads. None of them was judged as likely.

For the detection of large indels and CNV the approach based on the comparative coverage was used as described in section 5.3.3.3. The analysis using SSAHA2 output identified 11 such changes between the two strains. However, the visual analysis confirmed two of them: a 34bp deletion on chr 7 and a large 1.2 kb deletion on chr 5. They were also the only ones called by the analysis with MAQ software as well. The rest was classified as unlikely.

It has to be added here that indel identification algorithms are less stringent than those used for SNP identification. The simulation approach consisting on inserting artificial deletions into the genome “*in silico*” and replicating the SOLEXA re-sequencing showed that it is likely to generate a substantial amount of false positives however it is unlikely to miss the actual changes (Axel Martinelli personal communication).

	Chromosome	Type of change	Position	Size	comparative coverage	small indel quality index
Identified by SSAHA2						
1	1	Deletion	13,917	1		20/38
2	1	Deletion	19,767	1		28/50
3	1	Insertion	203,105	1		10/19
4	2	Insertion	32	1		2/3
5	3	Deletion	350,001	1		2/4
6	4	Deletion	104,874	1		14/16
7	5	Insertion	311,828	1		26/30
8	5	Deletion	410,773	1		2/3
9	7	Insertion	910,274	1		2/2
10	7	Deletion	910,327	2		2/3
11	7	Insertion	910,341	1		6/7
12	8	Insertion	1,184,534	1		2/3
13	9	Deletion	911,419	1		6/11
14	9	Deletion	1,301,305	3		13/25
15	10	Deletion	290,661	1		43/47
16	10	Deletion	718,122	1		57/71
17	11	Insertion	356,249	1		2/3
18	11	Deletion	760,507	1		24/24
19	11	Deletion	836,567	1		36/45
20	12	Insertion	142,723	1		4/8
21	12	Deletion	203,419	1		11/18
22	13	Deletion	872,812	1		5/8
23	13	Insertion	879,482	1		29/37
24	13	Deletion	1,188,222	1		21/30
25	13	Deletion	1,792,799	1		3/5
26	13	Insertion	2,115,249	1		5/9
27	14	Insertion	82,779	1		2/2
28	bin	Insertion	798	1		12/24
29	bin	Deletion	35,739	1		2/2
30	bin	Deletion	262,255	1		9/9
Identified by coverage analysis						
1	1	indel	581,324	75	0.02	
2	4	indel	793,932	50	0.17	
3	5	indel	683,724	1,266	0.24	
4	6	indel	372,860	24	0.00	
5	7	indel	169	13	0.00	
6	7	indel	13,193	17	0.18	
7	7	indel	876,927	34	0.19	
8	11	indel	788,143	17	0.23	
9	13	indel	2,307,750	12	0.25	
10	14	CNV	1,595,078	254	3.20	
11	14	indel	2,015,046	15	0.25	

Table 5.8 The list of indels identified by bioinformatic analysis.

a) represents the difference in comparative coverage between AS-sens and AS-30CQ (defined in section 5.3.3.3.). b) the quality index representing the relationship between the number of reads calling given indel (first number) and total number of reads covering the potential change (second number). The two deletions confirmed by MAQ and visual analysis are marked in green.

5.4.2. The validation of Solexa data

In order to confirm the results of the bioinformatic analysis a number of dideoxy-sequencing primers were designed. The results are presented in Table 5.9.

All seven high quality SNPs called with both software packages (see Table 5.6 and 5.7) were sequenced and confirmed. Also the smaller one out of the two predicted indels (table 5.8) could be confirmed (a 34bp deletion on chr 7). Unfortunately due to its size and position no successful primers could be designed for the larger 1.2 kb deletion on chr 5.

Additionally some of the changes identified with only one type of software were also sequenced. It included two putative SNPs on chr7 identified by MAQ software (see Table 5.6), a chr 13 SNP identified by SSAHA2 (Table 5.7) and two out of the 39 rejected deletions identified by SSAHA2 which were judged as most likely (covered by the high number of reads and called by over 90% of them). All of them proved to be false positives. It has to be mentioned here that attempts were made to sequence also the other four low quality SNP calls identified with SSAHA2 only (see Table 5.7) but no successful primers could be designed to amplify those,, probably because of their location in low complexity regions (three of them map to unassigned, telomeric contigs)

Finally, a number of SNPs rejected at the initial stages of bioinformatic analysis with MAQ software were also sequenced. These SNPs were situated within the selection valleys on chromosome 3 and 11 and predicted to cause the coding change within existing proteins, were also sequenced. None of those SNPs were confirmed.

In summary the dideoxy sequencing data showed 100% correlation with the results of bioinformatics analysis. All changes predicted as likely by bioinformatics analysis were present in the sequence and all potential mutations rejected because of the low call quality and/or lack of visual confirmation were absent.

Chromosome	Position	Called by	Type of change	Nucleotide change or indel size	Confirmed?
High quality calls by both types of software					
2	216,954	MAQ/SSAHA2	SNP	C/A	Yes
3	70,553	MAQ/SSAHA2	SNP	G/T	Yes
3	474,123	MAQ/SSAHA2	SNP	C/A	Yes
7	994,546	MAQ/SSAHA2	SNP	G/A	Yes
7	876,901	MAQ/SSAHA2	deletion	34bp	Yes
10	634,932	MAQ/SSAHA2	SNP	T/C	Yes
11	996,332	MAQ/SSAHA2	SNP	G/T	Yes
14	936,945	MAQ/SSAHA2	SNP	T/G	Yes
Low quality calls, rejected by visual inspection, identified only with one type of analysis					
13	113,002	SSAHA2	SNP	G/A	No
7	876917	MAQ	SNP	A/G	No
7	876919	MAQ	SNP	C/A	No
10	290,661	SSAHA2	Deletion	1bp	No
11	760,507	SSAHA2	Deletion	1bp	No
Initially identified by MAQ but rejected by later bioinformatic analysis					
3	180,102	none	SNP	G/A	No
3	180,086	none	SNP	A/C	No
11	946,436	none	SNP	G/T	No
11	946,392	none	SNP	C/T	No
11	432,447	none	SNP	T/C	No
11	1,013,413	none	SNP	C/G	No

Table 5.9 The validation of the analysis of Solexa results.

5.4.3. The summary of genetic changes between AS-sens and AS-30CQ

The Table 5.10 summarizes all of the predicted mutations occurring between AS-sens and AS-30CQ. In total seven point mutations and two deletions were proposed with high confidence and validated. Importantly, they include two previously identified point mutations. Those are PCHAS_072830 (*pcdhfr*) S106N, the main determinant of PYR resistance in the lineage (Cowman and Lew, 1990), and PCHAS_020720 (*ubp1*) V2728F that was identified in AS-30CQ in investigations regarding artemisinin resistance (Hunt et al., 2007). Now, 4 additional non-synonymous point mutations, 1 non coding point mutation and two deletions (one small, non-coding and one whole gene deletion) are revealed. All the mutations were identified in AS-30CQ only and absent in both AS-sens and in the reference sequence.

Chromosome	Position	Type of change	The base change (for SNPs) or size (for indels)	Gene ID	The change with the protein
2	216,954	SNP	C/A	PCHAS_020720	V2728F
3	70,553	SNP	G/T	PCHAS_030200	T707N
3	474,123	SNP	C/A	PCHAS_031370	T719N
5	683,724	Indel	1266 bp	PCHAS_051920	deleted
7	876,909	Indel	34 bp	intergenic	na
7	994,546	SNP	G/A	PCHAS_072830	S106N
10	634,932	SNP	T/C	PCHAS_101550	Y162H
11	996,332	SNP	G/T	PCHAS_112780	A173E
14	936,945	SNP	T/G	intergenic	na

Table 5.10 Predicted genetic changes between the AS-sens and AS-30CQ.

The previously known SNP are marked in green. The gene IDs are given according to the annotation from PlasmoDB from (September 2010 version)

5.4.4. The candidate mutations associated with CQ/ART resistance and their origin in the lineage

In this section I combine the genetic data from Chapter 4 with the results of the Solexa sequencing in order to determine the changes within predicted loci of resistance on chromosomes 11, 3 and 2 that may be connected with CQ/ART resistance.

The initial strategy to identify the origin of various mutations in the AS lineage was to sequence only the immediate progenitors of AS-30CQ, *i.e.* AS-sens, AS-pyr, AS-3CQ and AS-15CQ. However during these studies, concerns arose concerning the genotype of the AS-15CQ strain. It was the only uncloned strain in the lineage and, as a consequence, any genotyping was likely to be questionable. Therefore two cloned strains derived from AS-15CQ, namely AS-15MF and AS-ATN were also re-sequenced. As no further CQ selection was applied to any of these two branches of the lineage, any mutation connected to CQ resistance, found in AS-ATN, AS-15MF and AS-30CQ was likely to have its origins in AS-15CQ.

5.4.4.1. Chromosome 11

The LGS experiments with CQ selection characterised the most prominent selection valley to a central part of chromosome 11. This selection valley precisely coincides with the ~250 kb region between two specific AFLP markers defined to contain a locus conferring CQ resistance by previous classical linkage analysis (Hunt et al., 2004a). Within the entire chromosome 11 only one genetic change was identified - a G to T substitution at position 996,332 that is situated within the centre of the identified valley. It maps to the PCHAS_112780 gene (that I later refer to as *aat1* because of its putative function as amino acid transporter) and is predicted to cause a non-synonymous (A173E) substitution with the encoded protein..

The fragment of the *aat1* containing the predicted mutation was sequenced in all the strains from AS lineage and in AJ (Fig 5.1). The three CQ sensitive strains (AJ, AS-sens and AS-pyr) showed the wild-type G nucleotide and all of the CQ resistant strains starting from low resistant AS-3CQ have the mutated T nucleotide.

Therefore, as predicted, the mutation appeared during the selection of AS-3CQ from AS-pyr using low CQ concentrations. (Fig 4.17 in chapter 4).

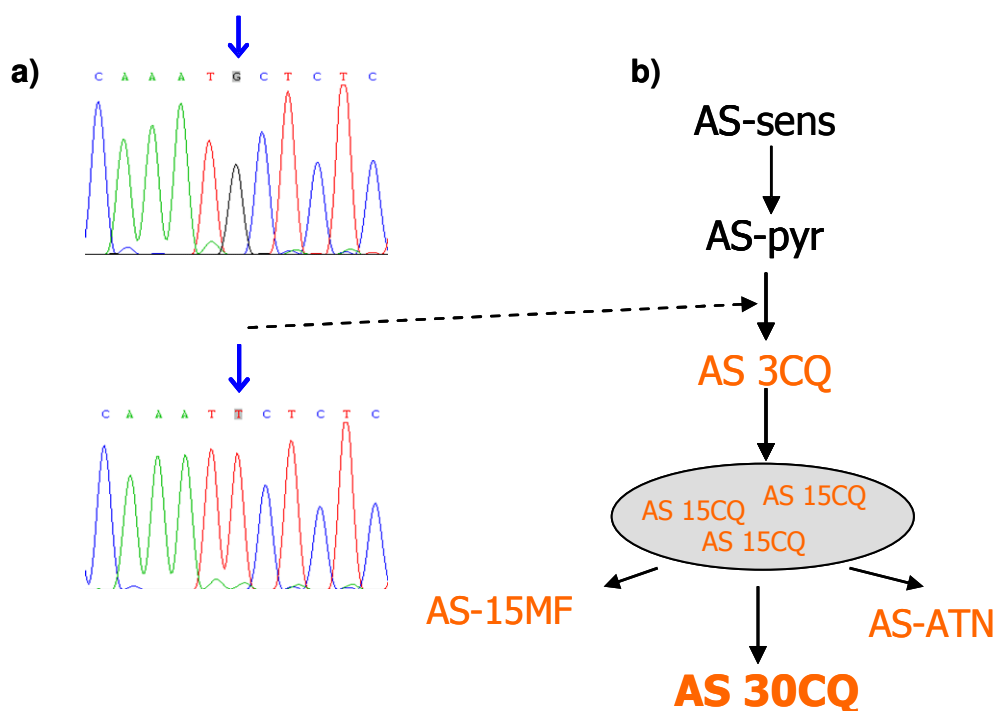


Figure 5.1 Point mutation resulting in A173E substitution in PCHAS_112780 (*aat1*) gene in various strains of AS lineage.

a) The chromatograms showing the substitution of the base in the *aat1* sequence b) the schematic of the AS lineage with the strains carrying the wt version of the *aat1* gene marked in black and those carrying the mutated version in red. The uncloned population of AS-15CQ strain is marked with a circle and its genotype is inferred from the genotyping of other strains in the lineage.

In summary, the *aat1* contains the only discovered mutation that arose within predicted locus and at the predicted step of selection and, therefore, is a very strong candidate for the main determinant of the CQ resistance in AS-lineage.

5.4.4.2. Chromosome 3

A secondary locus contributing to the CQ resistance was identified on chromosome 3 and Solexa genome re-sequencing identified two mutations here; G/T substitution at position 70,553 and C/A at position 474,123. They are both non-synonymous mutations in predicted genes. The C474123A mutation localises in a conserved *Plasmodium* spp. gene in the central syntenic part of the chromosome (PCHAS_031370, that was labelled *tm12* because of the presence of 12

transmembrane domains) and encodes for a T719N substitution. The G70553T translates into a T707N substitution in a *P.chabaudi*-specific gene (PCHAS_030200) in the non-syntenic part of the chromosome. Because the chr03 valley could not be further narrowed (*i.e.* all markers on chr03 were selected more or less equally), both mutations must be considered as possible candidates.

The sequencing of the two fragments of these genes in all strains of the lineage revealed that both mutations were absent in the sensitive (AJ, AS-sens, AS-pyr), low resistant (AS-3CQ) and intermediately resistant (AS-15CQ, AS-ATN and AS-15MF) strains. They were present only in the highly resistant AS-30CQ (Fig 5.2 and 5.3). These data suggest that either one or both of these mutations may confer highest level of CQ resistance which was in contrast with the previous predictions according to which the mutation on chromosome 3 was likely to be involved into the intermediate level of resistance and appeared between AS-3CQ and AS-15CQ.

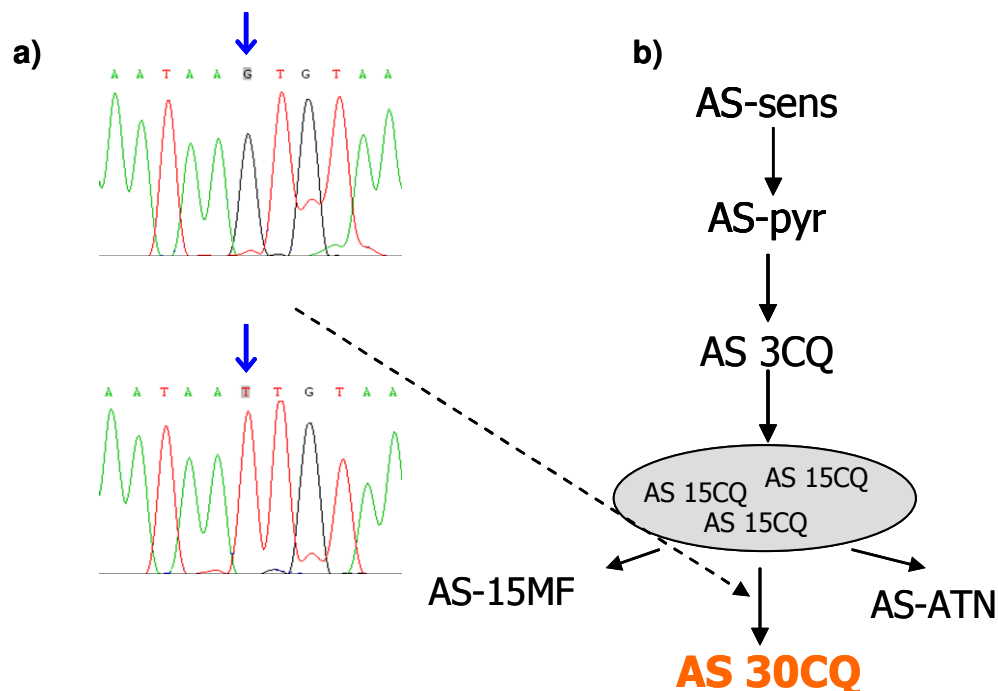


Figure 5.2 Point mutation resulting in T707N substitution in *PCHAS_030200* gene in various strains of AS lineage

a) The chromatograms show the substitution of the base in the sequence b) the schematic of the AS lineage with the strains carrying the wt version of the *PCHAS_030200* gene marked in black and those carrying the mutated version in red. The uncloned population AS-15CQ strain is marked with a circle and its genotype is inferred from the genotyping of other strains in the lineage. The dashed arrow points the place of most probable place of the appearance of the mutation within the lineage

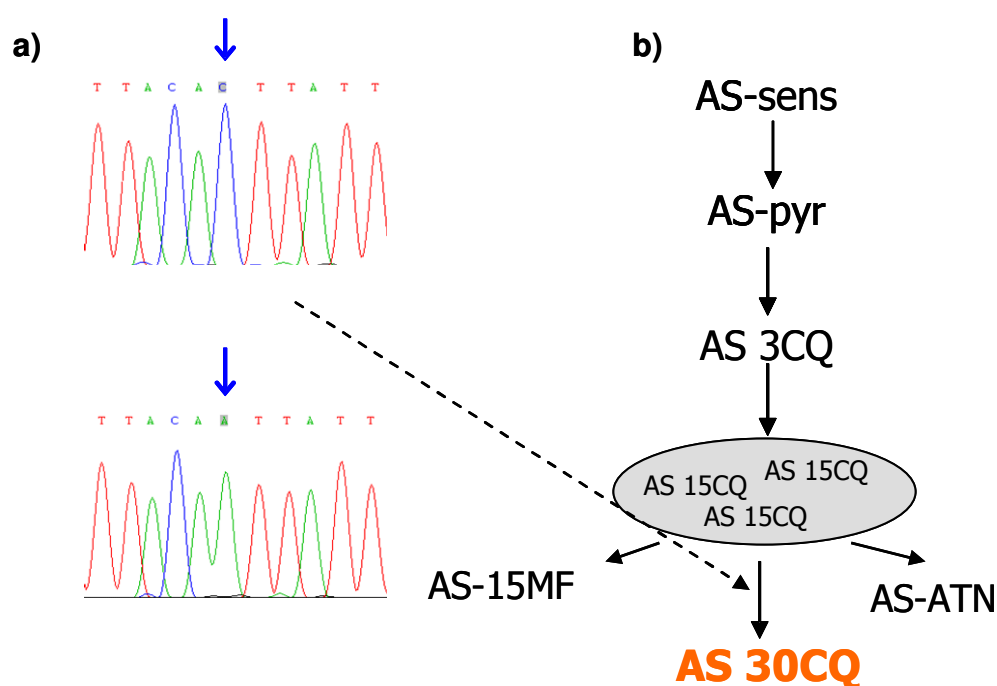


Figure 5.3 Point mutation resulting in T719N substitution in PCHAS_031370 (*tm12*) gene in various strains of AS lineage

a) The chromatograms show the substitution of the base in the sequence b) the schematic of the AS lineage with the strains carrying the wt version of the PCHAS_031370 gene marked in black and those carrying the mutated version in red.

During the investigation of AS-15MF from the AS lineage, however, another mutation was on chromosome 3 was discovered– a 3bp deletion (472273 – 5) (Sophia Borges, PhD thesis). Remarkably, it was also localised within *tm12* gene and encoded for a deletion of I102. Further investigation of the respective fragment of this gene in the AS lineage revealed that this mutation was present in all three intermediate resistant strains of the lineage – AS-15MF, AS-15CQ, AS-ATN), but absent in the highly resistant AS-30CQ (Fig .5.4). This is further discussed below in Section 5.5.2.

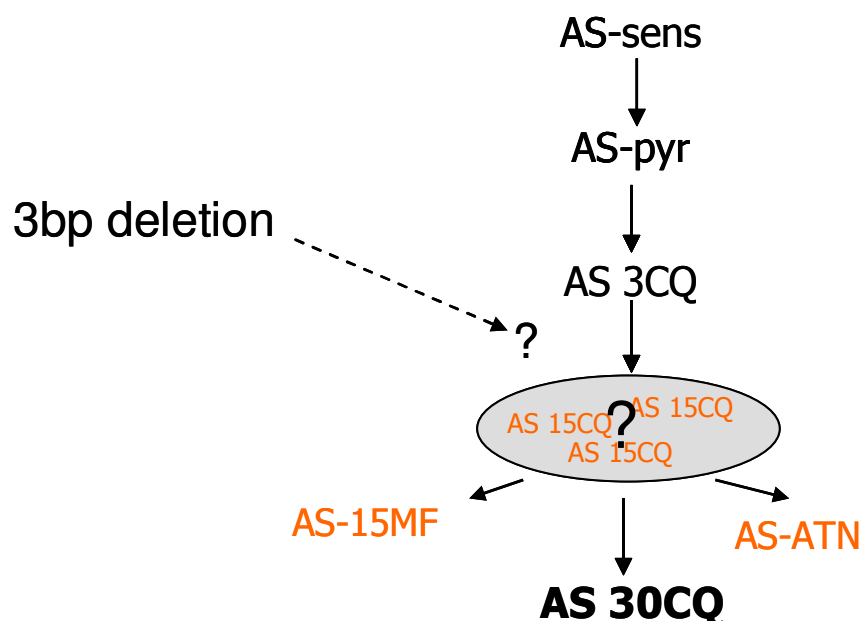


Figure 5.4 The 3bp deletion in PCHAS_031370 (*tm12*) gene in various strains of AS lineage. The AS lineage with the strains carrying the wt version of the *tm12* gene marked in black and those carrying the mutated version in red. The genotype of AS-15CQ strain and the place of appearance of the mutation within the lineage could not be established for the reasons mentioned in the text.

In summary, 2 potential determinants of CQ resistance were determined on chromosome 3 in AS-30CQ – PCHAS_030200 and PCHAS_031370 (*tm12*). The first one was mutated in highly CQ resistant strain only. The second one was mutated in both intermediate and highly resistant parasites, however different mutations were present the intermediate-resistant strain/clones AS-15CQ, AS-15MF, AS-ATN and the highly resistant clone AS-30CQ.

5.4.4.3. Chromosome 2

The selection of the AS-30CQ x AJ backcross with ART (Chapter 4, Fig. 4.15) confirmed the existence of the major locus on chromosome 2, as was the case with ART selection of the AS-ART x AJ cross (Hunt et al., 2007). Moreover a similar locus was selected with the highest CQ concentrations (Fig 4.7 and 4.8 d-e in Chapter 4). Previously, during the design of pyrosequencing markers, a mutation within the PCHAS_020720 (*ubp1*) gene was identified on this chromosome in AS-ART strain (a progeny of AS-30CQ) (Hunt et al., 2007). The connection between this mutation and ART resistance however, remained speculative because of the lack of the full inventory of mutations within the predicted region.

The Solexa sequencing results confirmed that, in AS-30CQ, only one mutation could be found on the whole chromosome 2 - the previously identified a C216954A substitution in PCHAS_020720 (*ubp1*) gene, that leads to V2728F mutation within the *UBP1* protein. Dideoxy sequencing (Fig. 5.5) revealed that three sensitive strains (AJ, AS-sens, AS-PYR1) and the low resistant AS-3CQ all bear the wild-type allele. The mutation was instead identified in AS-15CQ strain and two clones derived from it – AS-15MF and AS-30CQ. It would indicate that it did occur between AS-3CQ and AS-15CQ, coinciding with both an increase of CQ resistance the appearance of ART resistance. Surprisingly however, the third strain derived from AS-15CQ - AS-ATN - carries the wild type allele, which questions both the place of the appearance of the mutation and the AS-15CQ genotype (situation very similar to del I101 in *tm12*).

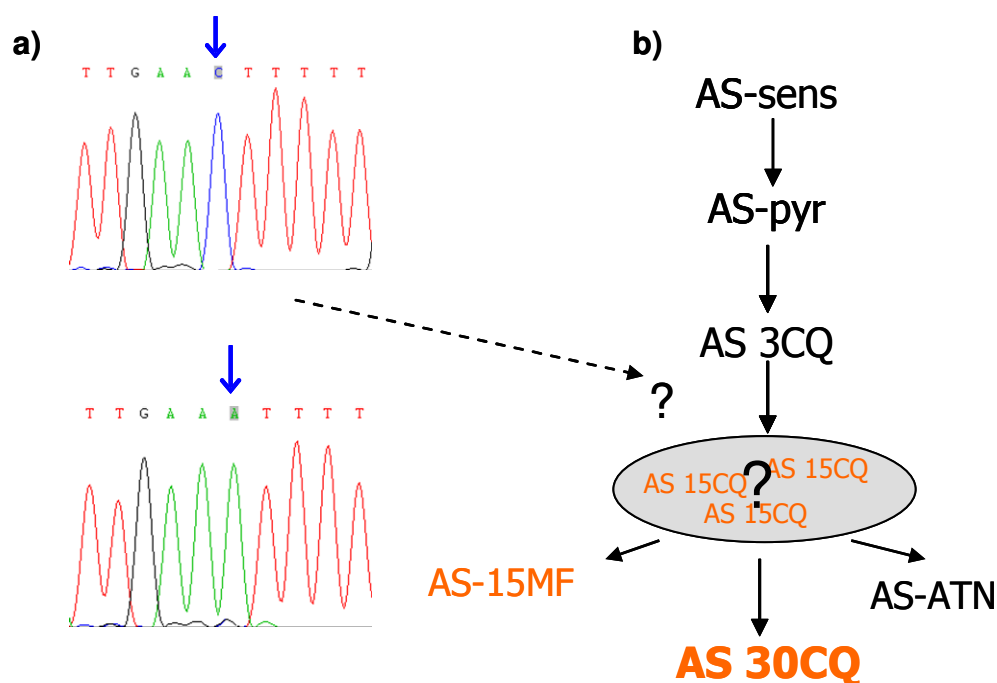


Figure 5.5 Point mutation resulting in V2729F substitution in PCHAS_020720 (*ubp1*) gene in various strains of AS lineage

a) The chromatograms show the substitution of the base in the sequence b) the schematic of the AS lineage with the strains carrying the wt version of the PCHAS_020720 (*ubp1*) gene marked in black and those carrying the mutated version in red

However, the sequencing of PCHAS_020720 in AS-ATN strain (Hunt et al., 2007) identified the alternative mutation within that gene, – a C217047A point mutation that results in V2697F substitution within the encoded protein (Fig 5.6).

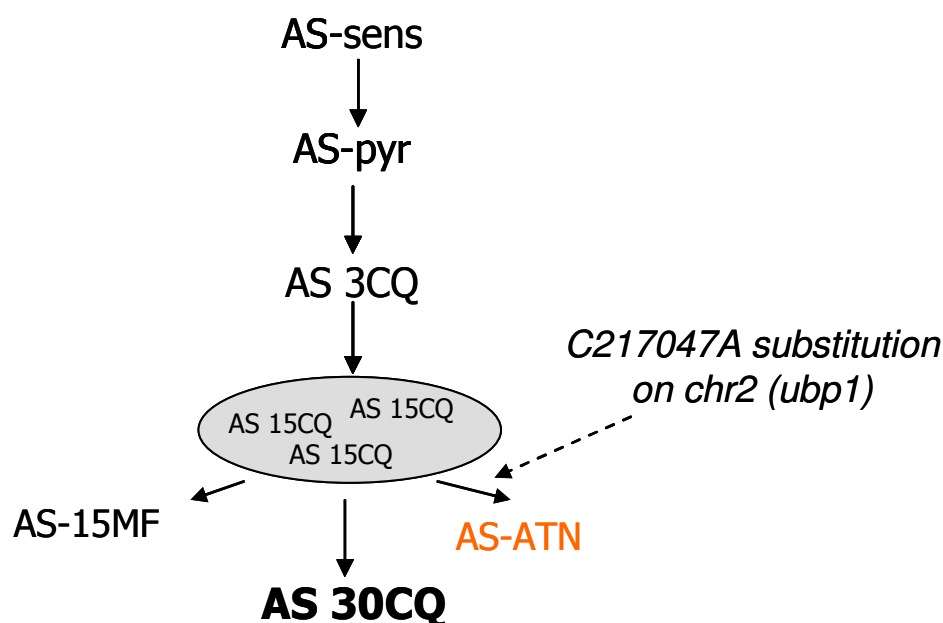


Figure 5.6 Results of previous sequencing of the point mutation resulting in V2697F substitution in *ubp1* gene in various strains of AS lineage.

The AS lineage with the strains carrying the wt version of the *ubp1* gene marked in black and the ones carrying the mutated version in red. The dashed arrow points the place of the most probable place of the appearance of the mutation within the lineage.

In summary, as no other genetic changes were present on chr 2, the *ubp1* mutation was very likely to be both the determinant of ART resistance and a contributing factor to the high CQ resistance in AS-30CQ strain. Additionally all the strains of the ART/CQ resistant strains (AS-15CQ, AS-30CQ, AS-ATN and AS-15MF) are carrying a mutated version of this gene. However, similarly to the *tm12* gene, two alternative mutations are present in the AS lineage – one in AS-ATN, another in the remaining clones. One possible interpretation of these findings is discussed in the section 5.5.2 of this chapter.

5.4.4.4. Other mutations identified by Solexa sequencing.

This work was performed in collaboration with Sofia Borges, Axel Martinelli and Louise Rodrigues

Beside the four mutations situated within the three major selection valleys that have just been described, five other genetic changes that arose between AS-sens and AS-30CQ were also investigated in the context of potential role in the CQ and ART resistance.

Three of them arose between AS-sens and AS-pyr, under pyrimethamine pressure and therefore were present in CQ and ART sensitive strain. They were:

- a previously known G994546A substitution on chr.7. It results in S106N mutation within PCHAS_072830 (*pcdhfr*) gene, encoding for dihydrofolate reductase that was previously identified as the main determinant of pyrimethamine resistance.
- the 34bp indel situated in intergenic region of chr 7
- the T936945G substitution on chr 14, in intergenic region.

The involvement of those mutations in CQ/ART resistance is very unlikely, as all of them were present in CQ- and ART-sensitive AS-pyr strain. They also fail to generate the repeatable selection valleys with either CQ or ART pressure during the experiments described in chapter 3 and 4 (one decreased marker on chromosome 7 observed under the highest ART concentration on Fig 4.16 is very likely to be an experimental error). Additionally one of them is a well-known mutation, responsible for pyrimethamine resistance in both human and rodent malaria and lacking any connection with CQ or ART resistance both in field and laboratory conditions. The other two are non-coding changes and therefore less likely to introduce any phenotypic change. Therefore they were no longer investigated in the context of resistance.

The two remaining mutations were:

- a 1.2 kb deletion on chromosome 5 that appeared causes a deletion of PCHAS_051920 gene (a possible antigen). This mutation couldn't be sequenced due to its size but Solexa sequencing of other strains in the lineage – AS-15MF and AS-ATN and AS-3CQ – (Paul Hunt, personal communication) strongly suggests that this mutation happened between AS-3CQ and AS-15CQ strains
- a T634932C substitution translating to Y162H mutation in PCHAS_101550 (a conserved Plasmodium protein of unknown function).

None of those mutations generate a strong signature in the experiments described in chapter 4 (although chr 5 is one of the many loci selected during the initial experiments with cross selection described in chapter 3 (see Fig. 3.4 and 4.9)).

However, they both appeared under CQ pressure and are introducing the coding mutations within the predicted proteins that may in some way impact the phenotype. Therefore, it cannot be excluded that they play some role in the studied resistance, but their influence either is not strong enough to generate a selection signature, or manifests itself in the different setting (e.g. during different treatment regime or in a different part of the life cycle).

5.4.4.5. Summary

The combination of the data generated in Chapter 4 with the Solexa sequencing suggests that at least four out of 9 mutations present in AS-30CQ may be connected with either CQ or ART resistance. More specifically our data suggests that *aat1* may be the main determinant of CQ resistance in our lineage and reinforces the hypothesis that *ubp1* is the main determinant of ART resistance. Additionally two possible gene candidates were located within a supplementary locus on chr 3 contributing to the high and intermediate level of CQ resistance (*tm12* and PCHAS_030200) and *ubp1* may be involved in the highest level CQ of resistance. Other mutations in the lineage failed to result in consistent deep selection valleys and therefore, either they are not involved in the studied resistance phenotypes or their effect is too small to be observed in our experiments.

The analysis of the appearance of these mutations throughout the AS lineage combined with the discovery of the alternative mutations within two out of four predicted resistance determinants (Table 5.11), however, reveals additional genetic complexity with the various pattern of mutations within *ubp1* and *tm12* genes observed in the intermediate and highly resistant strains. The possible interpretation of these data is discussed in section 5.5.2.

Strain	CQ Phenotype	ART phenotype	Genotype				
			Chr 11 <i>AatI</i>	Chr 2 <i>ubpI</i>	Chr 3 <i>tmI2</i>	PCHAS_030200	
AS-sens	sensitive	sensitive	173	2728	1102	719	707
AS-pyr	sensitive	sensitive	A	V	wt	T	T
AS-3CQ	Low resistant	sensitive	A	V	wt	T	T
AS-15CQ	Intermediate resistant	resistant	E	V	wt	T	T
AS-ATN	Intermediate resistant	resistant	E	F	deletion	T	T
AS-15MF	Intermediate resistant	resistant	E	V	deletion	T	T
AS-30CQ	Highly resistant	resistant	E	F	deletion	T	T
			E	F	wt	N	N

Table 5.11 The results of genotyping of CQ resistant lineage for mutations found on chr 11, chr 3 and chr 2

5.4.5. Proportional sequencing of predicted mutations in the selected backcross progeny

In order to confirm that the identified mutations may generate chloroquine and artemisinin selection signatures, the proportions of three of the four mutations identified as the likely resistance determinants (A173E in *aat1*, V2728F *ubp1* and T719N *tm12* were measured in selected backcross progeny generated by selection 2 during the experiments described in chapter 4 (see section 4.4.5).

The results are presented in Table 5.12 and on Fig. 5.7 and 5.8

Treatment	Percentage of <i>aat1</i> A173	Percentage of <i>tm12</i> T719	Percentage of <i>ubp1</i> V2728
No drug	75.0	74.4	91.1
1.5mg CQ	77.1	78.2	93.1
3mg CQ	1.0	11.6	80.7
10mg CQ	0.6	1.5	14.7
20mg CQ	0.2	0.7	8.6
100mg ART	38.2	35.4	8.8
200 mg ART	5.8	10.5	0.1

Table 5.12 The percentages of wt AJ allele within backcross progeny

5.4.5.1. A173E mutation in *aat1*

The proportions of wild type *aat1* allele in the untreated populations were ~75% and did not undergo any reduction at 1.5 mg CQ kg⁻¹ day⁻¹ which corresponded with the previous data reporting the lack of selection valley at this concentration, in this particular experiment.

However parasites bearing the wild-type allele almost disappeared during the other three CQ treatments – 3, 10 and 20 mg CQ kg⁻¹ day⁻¹. In all three cases the frequency of the wt allele in the population was below 1%, which is within an experimental error of the assay. Therefore, it is likely that nearly all parasites surviving the CQ treatments were carrying the mutated version of the *aat1* gene. The plotting of these values on the genome scan shows that this locus was the lowest point of the selection valley identified on chr 11 with all three concentrations (Fig. 5.7).

A distinct reduction in the frequency of the wild-type allele was also observed with ART treatments (particularly with 200 mg ART kg⁻¹ day⁻¹). However, this time the selection for the mutated allele is not complete (*i.e.* a significant portion of parasites surviving the ART treatment is carrying the wt *aat1* allele) and *aat1* is not the lowest point of chr 11. The potential explanation of selection on this chromosome can be found in section 4.5.4.

5.4.5.2. T719N mutation in *tm12*

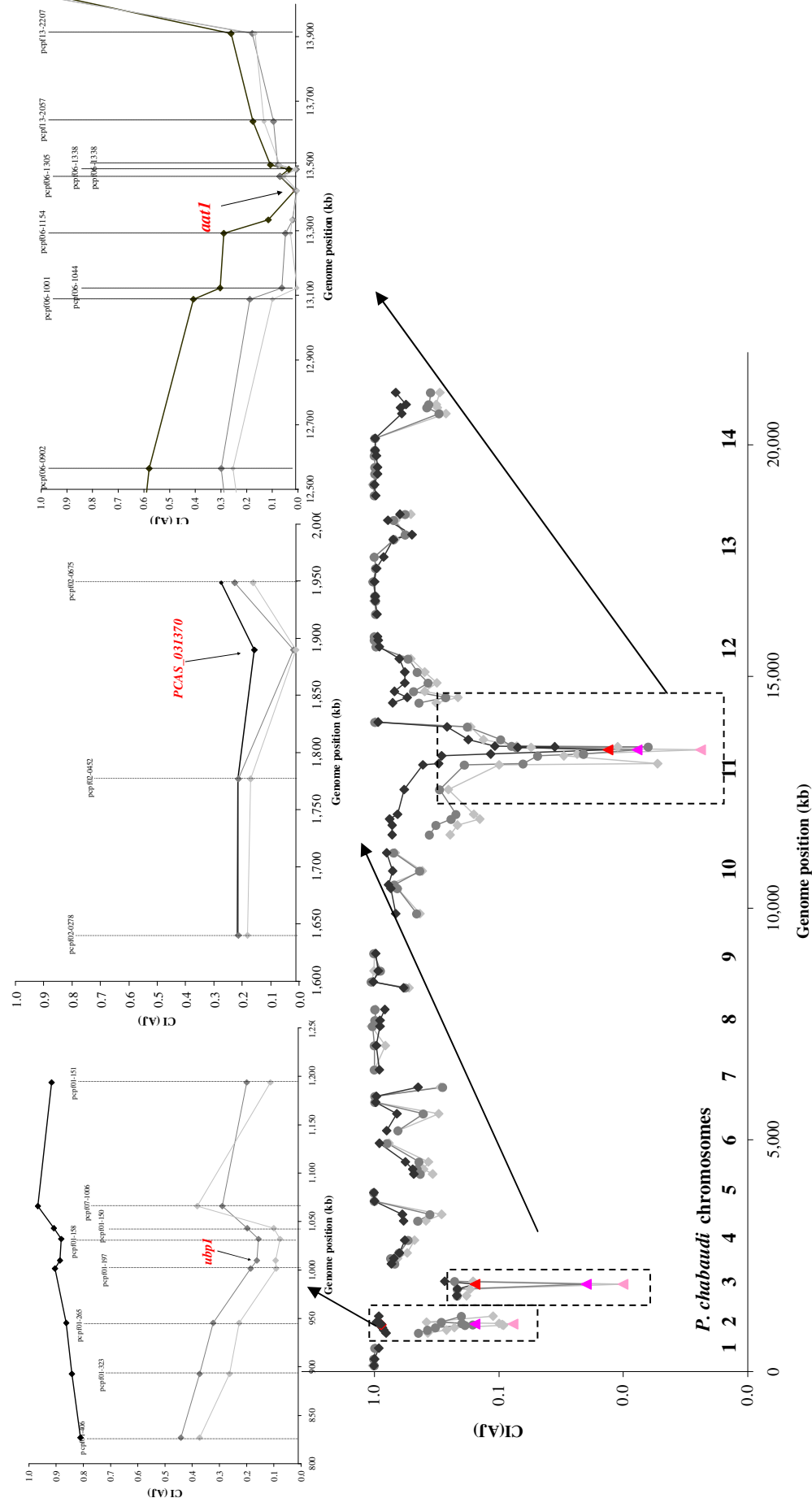
The wt allele is present in large proportions in untreated and in the 1.5 mg CQ kg⁻¹ day⁻¹ treated population. It is however strongly decreased in 3 mg CQ kg⁻¹ day⁻¹ treatment (which first shows evidence of a selection valley on chr 3) and completely eliminated with 10 and 20 mg CQ kg⁻¹ day⁻¹ treatment suggesting that it may be necessary to survive the high CQ doses. Again some evidence of selection can be observed under 200 mg of ART kg⁻¹ day⁻¹ however it can be easily explained by a significant scatter observed on this particular genome scan. In all selected populations this marker gives lower AJ allele values than any other marker on chromosome 3. These data therefore is consistent with the hypothesis that the *tm12* is the main factor contributing to further increase of CQ resistance after *aat1* mutation.

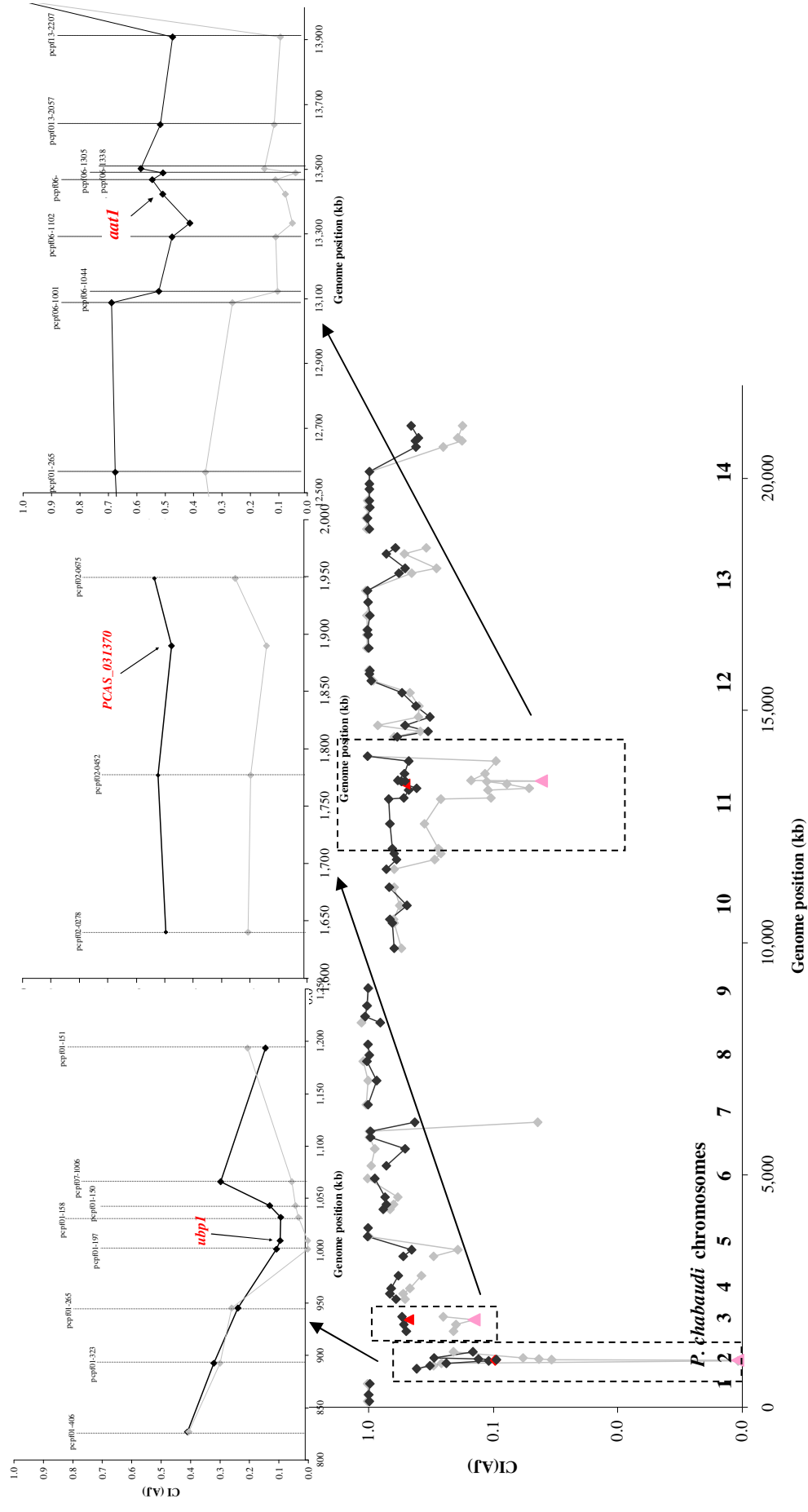
5.4.5.3. V2728F mutation in *ubp1*

In the populations grown without treatment or at 1.5 and 3 mg CQ kg⁻¹ day⁻¹, a very high (>80%) proportion of wt allele is observed. In the populations treated with 10 and 20 mg CQ kg⁻¹ day⁻¹ (those in which a selection valley on chr 2 was identified) a very high proportion of the mutant allele is observed instead. However in both cases the wt version of the gene is not completely eliminated in the population.

The strongest reduction of the wt version of the allele, however, is observed with ART treatment: at 100 mg/kg/day of ART the mutated allele is present in >90% of parasites and at 200 mg - in nearly 100%.

In all selected populations the proportions of AJ alleles of *ubp1* and/or the nearest pyrosequencing markers are the lowest of all the markers on chr 2.





5.4.6. The predicted determinants of CQ and ART resistance

This section briefly describes the four genes connected in CQ and ART resistance in *P.chabaudi*, summarizing the data concerning their function, expression, conservation in *Plasmodium* spp., the likely impact of the predicted mutations and how they might mediate the resistance phenotypes.

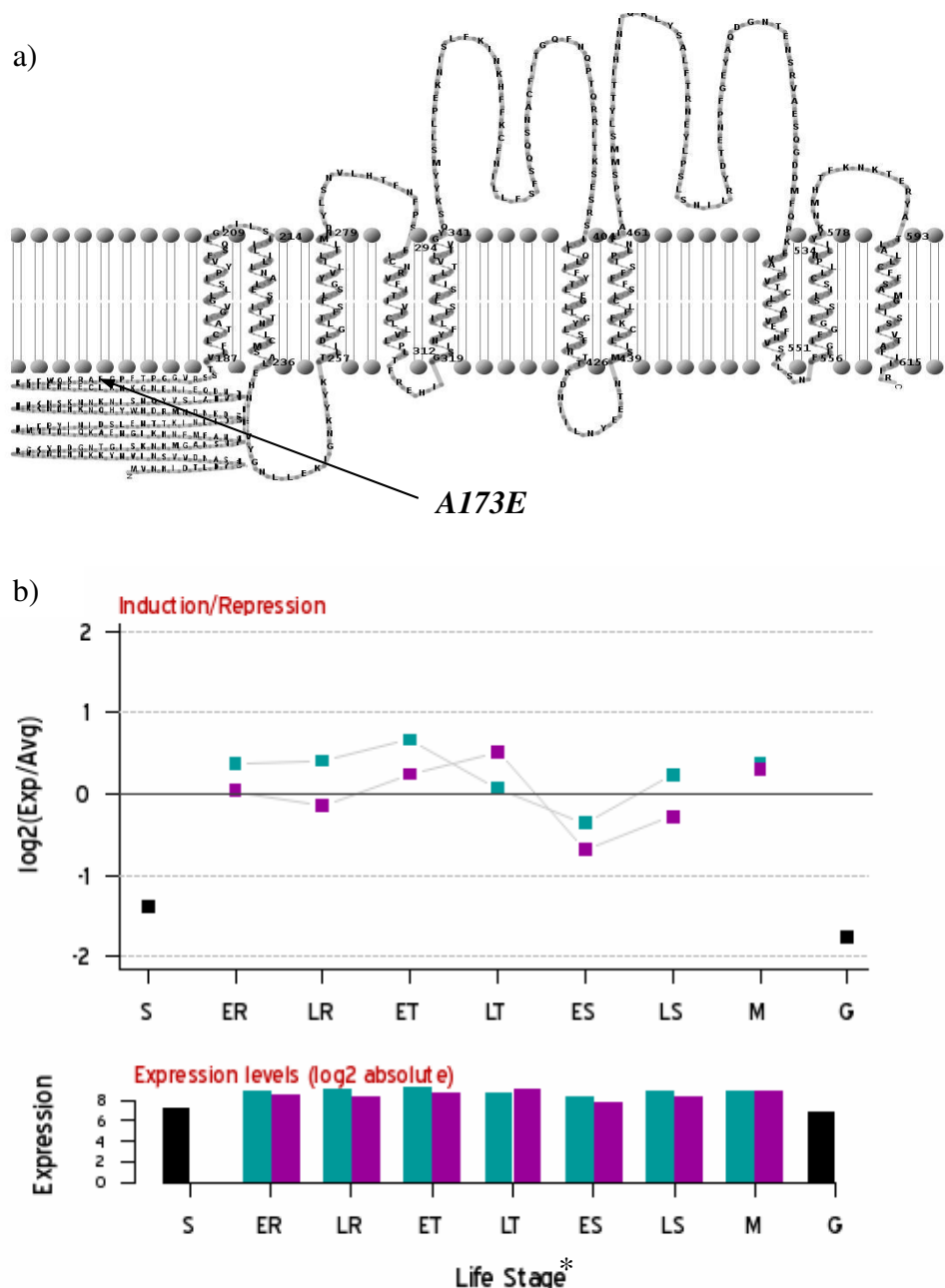
5.4.6.1.PCHAS_112780 (*Aat1*)

The main gene associated with CQ resistance, PCHAS_112780 (*aat1*), encodes a 70.6 kDa protein composed of 617 amino acids. Its predicted product is a transmembrane protein with 10 transmembrane domains (Fig. 5.9a).

No data is available concerning the expression of this gene in *P.chabaudi*. Its orthologue in *P.falciparum* (PFF1430c) however, is expressed mainly at the intraerythrocytic part of the cycle reaching its peak in the trophozoites (Fig. 5.9b). Similar data was obtained also in *P.vivax* orthologue (PVX_114575) (graph not shown). Therefore it seems likely that *aat1* follows similar expression profile reaching the peak of its expression at the stage targeted by chloroquine.

The protein sequence of *aat1* shows a high degree of conservation within *P.falciparum* species and within *Plasmodium* genus. Only 5 SNPs within this gene were discovered so far between various *Plasmodium* strains (see Table 5.13) and the alignment between *aat1* and its orthologues from three primate and two other rodent malarias shows that the core of the protein is practically identical in all six species (Fig. 5.10 and Appendix D).

What is the function of *AAT1*? Its orthologue in *P.falciparum* - PFF1430c - has been shown to contain a motif which defines its targeting to the digestive vacuole and this has been experimentally confirmed (D. Fidock and P. Moura, personal communication). Additionally PFF 1430c was recently identified as a potential member of the amino acid/auxin permease (AAP) family – a group of proteins responsible for transmembrane transport of amino acids in yeast, protozoans, plants and animals (Martin et al., 2005) . Based on sequence similarities Martin et al. predict that PFF1430c functions as a Na⁺ or H⁺ dependent general-specificity amino acid transporter. *AAT1* is likely to fulfil the similar role.



*Stages: ER = Early Rings LR = Late Rings ET = Early Trophs LT = Late Trophs ES = Early Schizonts LS = Late Schizonts
IM = Merozoites S = Sporozoites G = Gametocytes

Figure 5.9 The predicted structure of *AAT1* protein and the expression profile of its *P.falciparum* orthologue.

a) the predicted transmembrane structure of *AAT1* protein obtained using TMHMM software. The position of mutation discovered within *P.chabaudi* lineage is marked.

b) the expression of PFF1430c transcript in *P.falciparum* at various life cycle stages, obtained using Affymetrix oligonucleotide arrays according to (Le Roch et al., 2003). The erythrocytes cultures were synchronized either with sorbitol (cyan) or with temperature (violet). Sporozoite sample represents average of two replicates and gametocyte sample is available for sorbitol cultures only. The upper graph shows Log (base 2) ratio of Affymetrix MOID expression value (normalized by experiment) to average MOID value for all time points for a gene. The bottom one Affymetrix MOID expression value normalized by experiment. Graphs generated by PlasmoDB.

Position on chr	Position in protein	Major Allele	Minor Allele	Major Allele Frequency	Minor Allele Frequency	Phenotype	Major Product	Minor Product
1214261	null	C	G	0.71	0.29	non-coding		
1214352	552	G	T	0.67	0.33	syn	I	I
1215070	313	A	G	0.85	0.15	non-syn	F	S
1215235	258	G	A	0.82	0.18	non-syn	S	L
1215316	231	A	T	0.91	0.09	non-syn	V	D

Table 5.13 Polymorphisms in PFF1430c gene in *P.falciparum* population

SNPs obtained by combining the data Broad Institute, NIH and Wellcome Trust Sanger Institute (Jeffares et al., 2007; Mu et al., 2007; Volkman et al., 2007).

P.falciparum_PFF1430c	SN----DKKKKKKDSKK-----NSSNNNNNTTIVDI	139
P.vivax_PVX_114575	DGYYGEDNNNGAKDVQKNCQEGN-----ENENCDCDNQNSTIVDV	163
P.knowlesi_PKH_111940	DGYYDQDNNNDVNDTQENYQESN-----ATENYDNENENIAIVDV	189
P.chabaudi_PCHAS_112780	HWYHQNKNDNKKKNNKNSKNNKN-----ISNQYVSLAHVNDNDQEI	149
P.berghei_PBANKA_112830	NWYHQNKYNNKMNSKNSKN-----SHNKHISIAHIKYDDQEI	146
P.yoelii_PY04168	NWYHQNKYNNKMNSKNSKNSKNSKSSKNSKSHNKHISIDHIKDDQEI	158
	. .: .: .:	
P.falciparum_PFF1430c	SDGD-YTNDEEGTNKPKRNWKGRTFSRFTPGGVRSSSTVLFICTAIGVGFL	188
P.vivax_PVX_114575	GENEKNPEDEEAES-KKNWQRRRTFSRFTPGGVRSSSTVLFICTAIGVGFL	212
P.knowlesi_PKH_111940	GGNEKNAQDKEEEKKKKNWKRRTFSRFTPGGVRSSSTVLFICTAIGVGFL	239
P.chabaudi_PCHAS_112780	NENGKKNLCEENETNTKKTWQKRRTFGPFTPGGVRSSSTVLFICTAIGVGLL	199
P.berghei_PBANKA_112830	NEKGKNLYEENQNTTKTWKRRTFSRFTPGGVRSSSTVLFICTAIGVGLL	196
P.yoelii_PY04168	NEKEKNKIYEENQNTKKTWKRRTFSRFTPGGVRSSSTVLFICTAIGVGLL	208
	. : * . . * : . : * : * . : * : * : * : * : * : * : *	
P.falciparum_PFF1430c	SIPYVFSKLGIIILSIILILNFAESYVTNLTSSLEHNTFVYGNLLKK	238
P.vivax_PVX_114575	SIPYVFSKLGIIILSLLLILNFAESYVTNLTSSLEHNTFVYGNLLKK	262
P.knowlesi_PKH_111940	SIPYVFLKLGIIILSLLLILNFAESYVTNLTSSLEHNTFVYGNLLKK	289
P.chabaudi_PCHAS_112780	SIPYVFSKLGIIILSIILILNFAESYITNLTSSLEHNTFVYGNLLKK	249
P.berghei_PBANKA_112830	SIPYVFSKLGIIILSIILILNFAESYITNLTSSLEHNTFVYGNLLKK	246
P.yoelii_PY04168	SIPYVFSKLGIIILSIILILNFAESYITNLTSSLEHNTFVYGNLLKK	258
	***** : ***** : * : * : * : * : * : * : * : * : *	

Figure 5.10 The fragment of alignment of homologues of AAT1 protein between six Plasmodium species.

The alignment was produced with ClusatalW2 software. The position of the residue mutated in AS-30CQ is highlighted. The full alignment is available in appendix D.

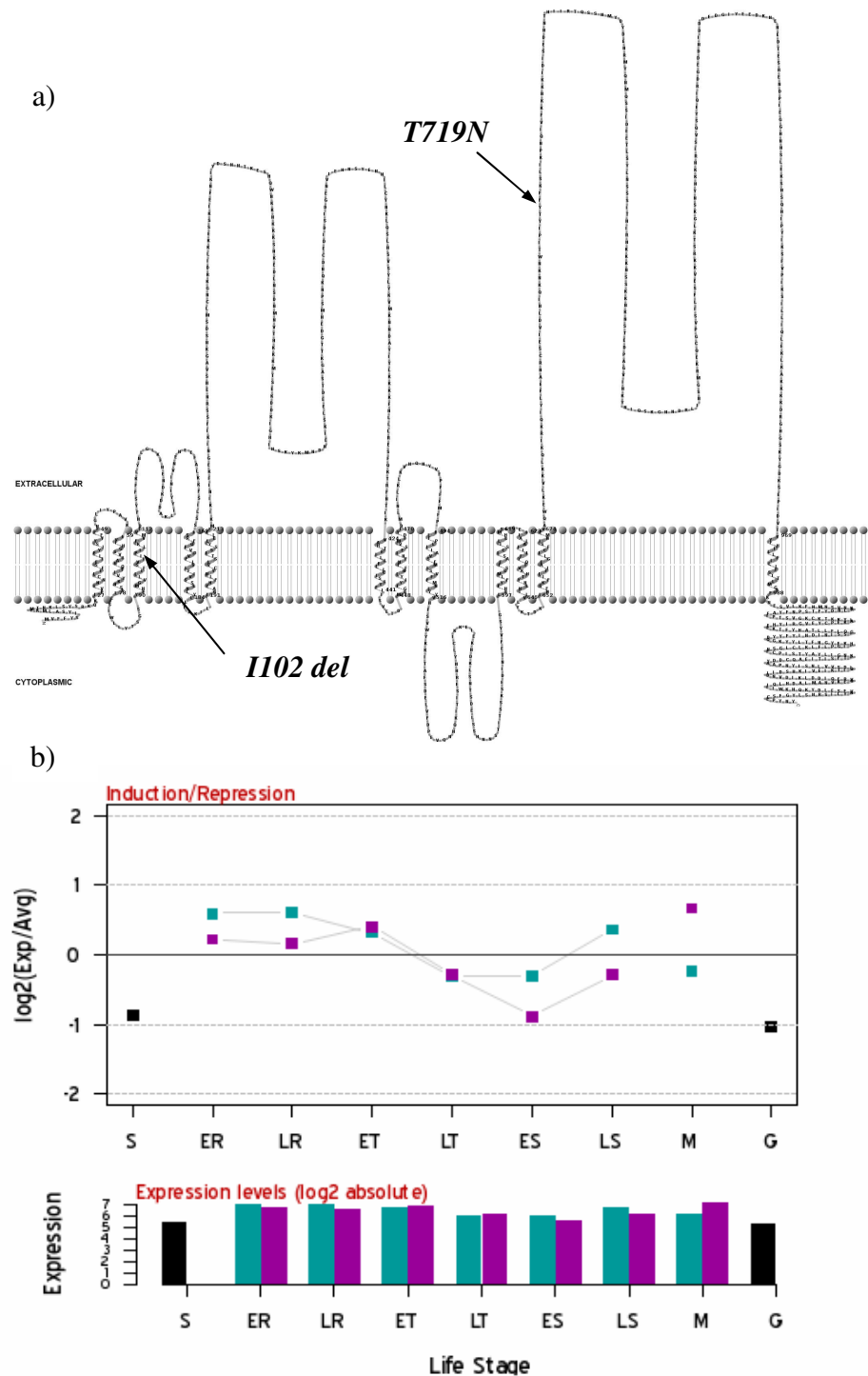
The mutation associated with the CQ resistance in our model is causing a change from alanine in position 173 into glutamic acid. This change is localised just before the 1st predicted transmembrane domain of the protein, on the “outer” side of the membrane (equivalent to the interior of digestive vacuole) (Fig.5.9a). It maps to the beginning of a highly conserved region and concerns a semi-conserved residue - the alignment between six Plasmodium species reveals that either alanine (A) (*P.chabaudi* and *P.berghei*) or threonine (T) (*P.falciparum*, *P.vivax*, *P.knowlesi* and *P.yoelii*) is present on the homologous position. Both A and T are small, uncharged residues. The mutation present in AS-30CQ instead introduces at this position more bulky, charged residue (E), and therefore is likely to affect the protein function (*e.g.* change the transporter specificity).

5.4.6.2.PCHAS_031370 (*tm12*)

The *tm12* gene was identified as one of two possible determinants of high or intermediate CQ resistance localised on chr 3. It is predicted to encode a 142.5 kDa, 1218 aa, hypothetical protein. Structure predictions suggest that it may be another membrane protein, containing 12 transmembrane domains (Fig. 5.11a).

Its homologue in *P.falciparum* (PFB0675w) and *P.vivax* (PVX_002795) are expressed mainly in merozoites and rings with the minimum of expression on schizonts (Fig 5.11b). The mass spectroscopy however, confirmed its presence in trophozoites (main targets of CQ action). The PFB0675w protein, similarly to the previous one, is predicted to target the digestive vacuole membrane (D. Fidock and P. Moura, personal communication). Its function remains unknown.

Also this gene is characterised by high degree of conservation. Only 15 SNPs were identified within *P.falciparum* strains (Table 5.14) and the alignment between different Plasmodium species reveals several regions highly conserved across all six parasites (Fig 5.12 and Appendix D).



*Stages: ER = Early Rings LR = Late Rings ET = Early Trophs LT = Late Trophs ES = Early Schizonts LS = Late Schizonts
IM = Merozoites S = Sporozoites G = Gametocytes

Figure 5.11 The predicted structure of *TM12* protein and the expression profile of its *P.falciparum* homologue

a) the predicted transmembrane structure of TM12 protein obtained using TMHMM software. The position of mutations discovered within *P.chabaudi* lineage is marked.

b) the expression of PFB675w transcript in *P.falciparum* at various life cycle stages obtained using Affymetrix oligonucleotide arrays according to (le Roch et al., 2003). The expression data presented in the same way as on Fig. 5.9.b)

Position on chr	Position in protein	Major Allele	Minor Allele	Major Allele Frequency	Minor Allele Frequency	Phenotype	Major Product	Minor Product
605144	34	T	G	0.89	0.11	non-syn	C	G
605351	null	T	G	0.88	0.13	non-coding		
605412	68	C	A, T	0.7	0.3	non-syn	A	E, V
605419	70	A	T	0.9	0.1	non-syn	K	N
605425	72	A	T	0.9	0.1	non-syn	K	N
605645	146	A	T	0.71	0.29	non-syn	N	Y
606435	409	G	A	0.91	0.09	non-syn	G	D
606666	486	G	A	0.93	0.07	non-syn	R	H
607100	631	G	T	0.92	0.08	non-syn	A	S
607784	859	A	C	0.95	0.05	non-syn	K	Q
608130	974	A	G	0.92	0.08	non-syn	D	G
608255	1016	A	C	0.64	0.36	non-syn	N	H
608501	1098	G	T	0.77	0.23	non-syn	A	S
608550	1114	A	C	0.77	0.23	non-syn	E	A
608790	1194	A	G	0.5	0.5	non-syn	H	R

Table 5.14 Polymorphisms in PFB0675w gene in *P.falciparum* population

SNPs obtained by combining the data Broad Institute, NIH, and Wellcome Trust Sanger Institute (Jeffares et al., 2007; Mu et al., 2007; Volkman et al., 2007)

Two mutations appeared in within *TM12* protein in AS lineage: I102del and T719N. The I102 is located within a central part of highly conserved sequence of hydrophobic residues forming the core of TMD 3 and is conserved throughout all six *Plasmodium* species (Fig 5.11a and 5.12). Additionally, the *in silico* prediction of secondary structure of wild-type and mutated protein suggests that the deletion of this amino acid is causing a significant rearrangements within the protein e.g. 10 instead 12 transmembrane domains are predicted. (Fig. 5.13). All this indicates that this residue may play a crucial role in the maintenance of correct protein structure, and its removal is likely to disturb the protein function.

The second change (T719N) is an amino acid substitution that takes place on the "outer" side of the membrane (DV inner), in the middle of the large loop of the protein between transmembrane domains 11 and 12 (Fig 5.11a). This residue is semi conserved throughout the *Plasmodium* genus. The homologous position contains threonine in all rodent parasites and serine in all primate ones. It suggests the importance of a hydroxyl-containing side-group on this position e.g. as potential phosphorylation or binding site, which is removed by the discovered mutation.

```

P.falciparum_PFB0675w      NGIINPAVKIVFHGSIIEDKTKVYENSIIESIWLFLTSGSYFNMAMLLFF 100
P.vivax_PVX_002795        NGIINPAVKIVFHGNIADNTQVYENSILESIWLFLTSGSYFNMAMLLFF 100
P.knowlesi_PKH_040510     NGIINPAVKIVFHGNIADTTQVYENSILESIWLFLTSGSYFNMGMLLFF 99
P.chabaudi_PCHAS_031370   NGLIYPAVKIVFHGSVIANQTIVYKNSILESIWLLYTSGSYFNMGVLLIF 99
P.berghei_PBANKA_031160   NGLIYPAVKIVFHGSVIANQTIVYKNSILESIWLLYASGSYFNMGVLLIF 99
P.yoelii_PY05194         -----MGVLLIF 7
                             *.:**.*

P.falciparum_PFB0675w      SIIPFLKLLMVSDNFYSFIVLYKMNNKREEEKRERRRRIRAREYNNAKY 150
P.vivax_PVX_002795        SIIPILKLIMVSDNFYCFCKLYNLSLQHEEEEEKEE----- 137
P.knowlesi_PKH_040510     SIIPILKFIMVSDNFYCFCKLYNLSLQHEEEEEKEE----- 136
P.chabaudi_PCHAS_031370   SVIPILKFVMVTDNFYSFFKLYNLGLKFENENDKDD----- 135
P.berghei_PBANKA_031160   SVIPVLKFVMVSDNFYSFFKLYNLSLNFENENDKDD----- 135
P.yoelii_PY05194         SVIPVLKFIMVSDNFYSFFKLYNLSLNFENENDKDD----- 43
                             *:***.***:***.***.*: **:. : :*:

P.falciparum_PFB0675w      NKNRYKMKNYNRNKNINDIYKDDIYYSENIFKNDEINYINTINENEFFIL 200
P.vivax_PVX_002795        -----EDDIFFINTVNENEQFIL 155
P.knowlesi_PKH_040510     -----EDDIFFINTVNENEQFIL 154
P.chabaudi_PCHAS_031370   -----AKLIYYINIVNENEFFAL 153
P.berghei_PBANKA_031160   -----AKLIYYINIVNVNGEFAL 153
P.yoelii_PY05194         -----AKLIYYINIINENEFFAL 61
                             . * : ** : * * : * *

P.falciparum_PFB0675w      KSEKIKFCKFLFLFILYGLLCSINLYVNQWEEYITKLKRLKRRILLFEKN 825
P.vivax_PVX_002795        KSNTVKFCKFLFLFVYGLLCLINLYVSQWELYVKKLKKTKKRIALFQNN 776
P.knowlesi_PKH_040510     KSSTVKFCKFLFLFVYGLLCLINLYVSQWELYVKKLKKTKKRIALFQNN 776
P.chabaudi_PCHAS_031370   KTDKIHFCKFILFISFGLICFSINMYVAQWEKLVKKLQKIRKQIYLFESN 697
P.berghei_PBANKA_031160   KTDKIHFCKFILFISFGLICFSINMYVTQWEKIIKKLQKIRKQIFLFESN 693
P.yoelii_PY05194         KTDKIHFCKFILFISFGLICFSINMYVAQWEKLIKKLQKIRKQIFLFESN 594
                             *:..:*.***: **:.***:*** ** * :.***: :*: * * : *

P.falciparum_PFB0675w      KFSEIVDLNTQKGDGDFDETQIFSIFFSFLIKKNE--GSKMRDNDMNSD 873
P.vivax_PVX_002795        RYSEIVDLNIQK---HEHWHVPIYSYFFQFLKKKEHFFSSSLKHNHITK- 822
P.knowlesi_PKH_040510     RYSEIVDLNIQK---HEHWHVPIYSYFFQFLKKKEHFFSPSIKHNQITK- 822
P.chabaudi_PCHAS_031370   RCSEIVDLNIQK---NKNWTIPIYSYFFKSLIKPNQIVFKKNNKIDISE- 743
P.berghei_PBANKA_031160   RCSEIVDLNIQK---NKNWNIPIYSYFFKFLIKQNQNILKKKKIDISQN 740
P.yoelii_PY05194         RCSEIVDLNIQK---NKNWNIPIYSYFFKFLIKQNQNIFKKKNKINISQD 641
                             : ***** ** .. *: **.*:* .. .. :..

P.falciparum_PFB0675w      SEDSIYDAYEQQIQHLHGDNMVNGMLMMRRISMQNLEDDDETQ----- 915
P.vivax_PVX_002795        --EHTDTVQTMMHVRVYENDTSLGYQNTRVDNDELNDFNEFN----- 862
P.knowlesi_PKH_040510     --EHTDTVQTMMYRVYENDTSLGYQNTRVDND---EMVEFN----- 859
P.chabaudi_PCHAS_031370   --DNNTHIIRTQSSNMYDYDENR----- 764
P.berghei_PBANKA_031160   INNN--HIIRTQSSMTYDNDENK----- 761
P.yoelii_PY05194         INNNNNHIIRTQTSTTYDCDANQNYQNCQNYQNYQNCQNYQNCQNYQNCQ 691
                             : : :

```

Figure 5.12 The alignment of two fragments of *TM12* protein between six *Plasmodium* species. The positions of I102del and T719N mutations are highlighted. The full alignment can be found in the Appendix D.

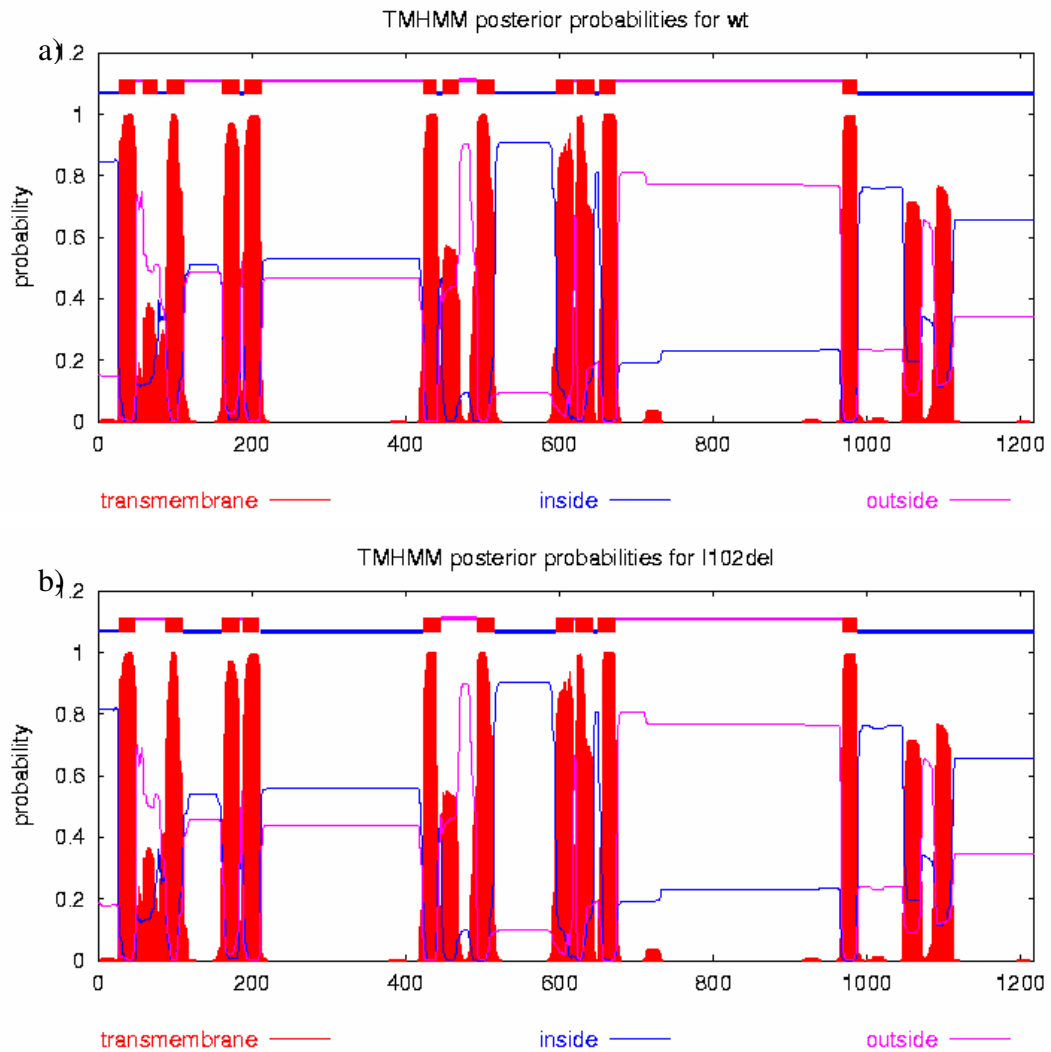


Figure 5.13 The impact of I102del in *TM12* on protein structure

a) the pattern of transmembrane domains predicted for wt protein b) the pattern of TMD predicted for a mutant. The prediction for the mutant protein is more likely to reflect the constraints in assuming the correct structure than the actual, new conformation.

5.4.6.3.PCHAS_030200

A second one of the two proteins mutated on chromosome 3. Because this gene is non-syntenic, *i.e.* localised within the *P.chabaudi*-specific part of the genome, the information about its structure and function is very limited. It encodes for a hypothetical 1287 amino acid protein belonging to the interspersed repeats (*pir*) superfamily. The members of that family in *P.chabaudi* are located in telomeric regions and expressed mainly at trophozoite stage (Janssen et al., 2004). At least some of the members of this family were shown to be exported from the parasite and located at the surface of infected erythrocyte (Sharling et al., 2007). Their exact function is not known - immune evasion, erythrocyte invasion, intracellular transport and intracellular signalling have been suggested (Cunningham et al., 2010).

Due to the lack of data, no hypothesis can be formulated about the potential role of the T707N substitution that appeared in this protein under high CQ selection.

5.4.6.4.PCHAS_020270 (*Ubp1*)

As the only mutated gene on chromosome 2, *ubp1* was found to be both the likely main determinant of ART resistance and the contributing factor to the high CQ resistance. This gene encodes a large 416.1 kDa protein composed of 2921 amino acids. Based on the data from *P.falciparum* and *P.vivax* orthologues, it is expressed throughout the lifecycle with the peak of expression in merozoites and rings and minimal expression in early schizonts (Fig. 5.14b).

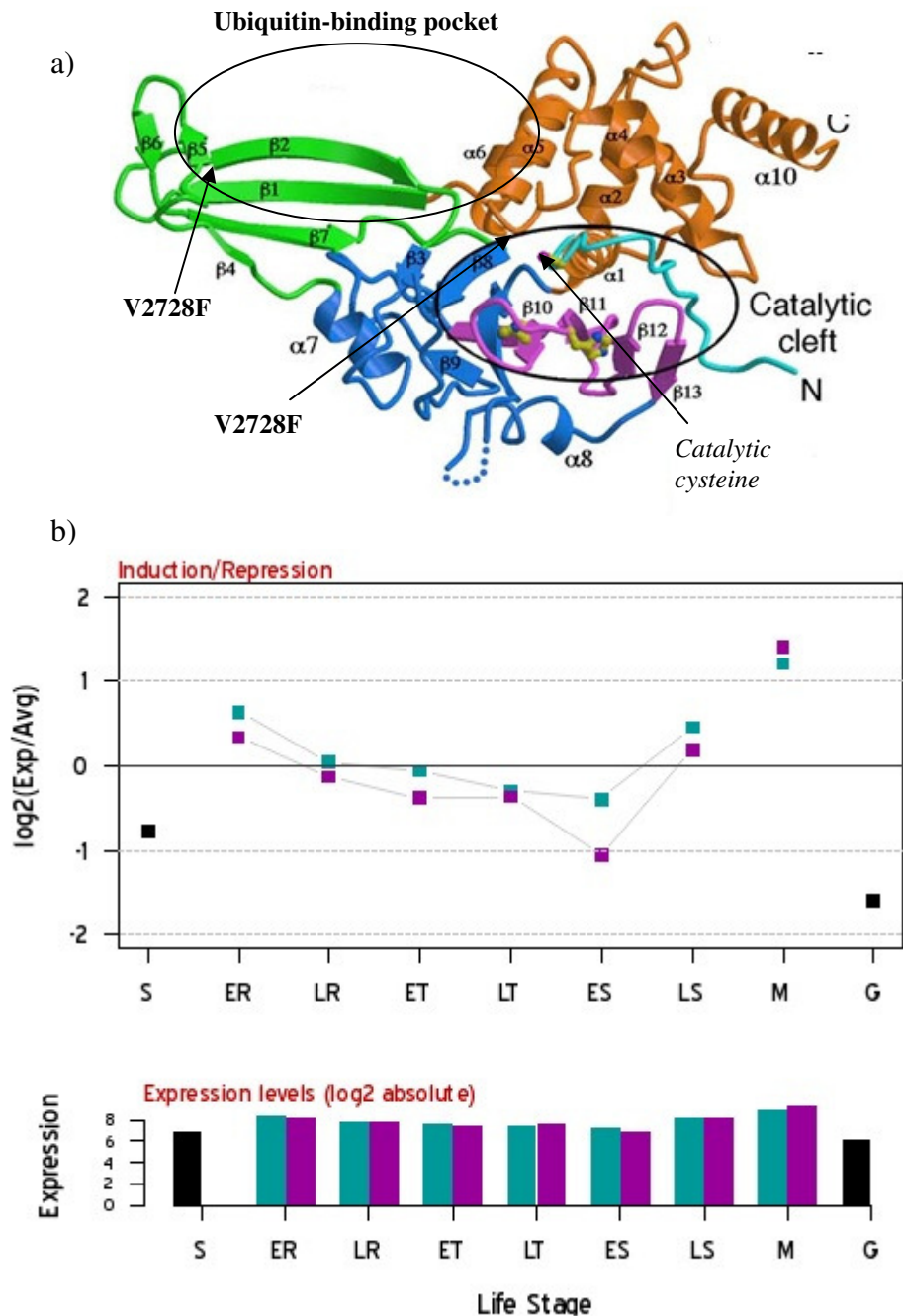
Its predicted function is a ubiquitin carboxyl-terminal hydrolase, a de-ubiquitinating enzyme. No molecular targets of this enzyme have been identified so far, although its *P.falciparum* homologue (PFA0220w) was shown to interact with four proteins in the budding yeast experimental system (LaCount et al., 2005). Those are: PF11_0477 (CCAAT-box DNA binding protein), PF13_0050 (HORMA domain protein, potentially involved in mitosis), MAL7P1.155 (a zinc finger protein) and MAL8P1.153 (a transcription factor). This preliminary results have to be treated with caution, however they suggests the potential involvement in transcription regulation and/or cell division.

In this case the protein is not as highly conserved as *AAT1* and *TM12*. The number of polymorphisms present in *P.falciparum* gene is much higher (56, see Table 5.15) and the alignment between six *Plasmodium* species shows less than 50% of identical residues between the species (see Appendix D). However, it also identifies a highly conserved region at the C-terminus of this protein (the predicted catalytic domain), which is almost identical in all sequenced species and contains no coding polymorphisms in *P.falciparum* population. Both mutations identified in AS lineage (V2697F and V2728F) are located within this region and change the residues conserved throughout the *Plasmodium* genus (Fig. 5.14a and 5.15).

The study of the possible effects of the V2728F and V2697F mutations on this protein was published previously (Hunt et al., 2007). In summary, the conserved C-terminal part of the *UBP1* protein was used to interrogate the protein structure database revealing the similarity to the human herpes-virus-associated, ubiquitin-specific protease (HAUSP) for which the crystal structure has been solved (Hu et al., 2002). This structure was used as a template for a 3D model of *UBP1* catalytic domain (presented on Fig 5.14a), revealing the position and potential role of the mutated residues. It indicates that both mutations are likely to cause a significant shift in protein structure, potentially affecting its function.

The V2728 residue (equivalent of HAUSP I332) in the wt *UBP1* protein points to a hydrophobic pocket required for ubiquitin binding, defined principally by Y2779 (UBP-1) and F4 (ubiquitin) (Fig. 5.14a and 5.16a). The introduction into this spot of a much more bulky phenylalanine is causing dramatic clashes in the side chain orientation and the whole structure of the protein needs to be modified in order to accommodate this change. Therefore the disruption of ubiquitin binding and the decrease in the enzymatic activity of the protein is very likely.

The AS-ATN specific V2697F mutation (equivalent of HAUSP V296) maps to an end of alpha helix 5 close to the proposed catalytic cysteine, C2586 (Fig. 5.14a). A wt residue fits in the hydrophobic pocket defined by Y2857 and I2845. Again the predicted mutation introduces the bulky residue that forces a significant rearrangement of the neighbouring amino acids. It is most likely that the phenylalanine side chain would point into a small pocket produced by V2905, I2845 and L2591 (Fig. 5.16.b). However, the V2697F side chain comes too close to these residues to be accommodated without further movement. These movements are likely to affect the catalytic cysteine, C2586, found on the same helix as L2591. It is also likely that ubiquitin binding will be affected, because Y2857 interacts directly with ubiquitin. Therefore also this mutation is predicted to interfere with the catalytic function of the protein.



*Stages: ER = Early Rings LR = Late Rings ET = Early Trophs LT = Late Trophs ES = Early Schizonts LS = Late Schizonts
IM = Merozoites S = Sporozoites G = Gametocytes

Figure 5.14 The predicted structure of the fragment of *UBP1* protein and the expression profile of its *P.falciparum* homologue.

a) the predicted 3D structure of the conserved, C-terminal domain of *UBP1* protein. The positions of the mutations discovered within *P.chabaudi* lineage, as well as the place of ubiquitin-binding side and the catalytic cleft with the crucial cysteine residue, are marked.

b) the expression of PFA0220w transcript in *P.falciparum* at various life cycle stages obtained using Affymetrix oligonucleotide arrays according to le Roch et al., 2003. The expression data is presented in the same way as on Fig. 5.9.b.

Position on chr	Position in protein	Major Allele	Minor Allele	Major Allele Frequency	Minor Allele Frequency	Phenotype	Major Product	Minor Product
190710	68	A	G	0.92	0.08	syn	L	L
190712	69	A	G	0.92	0.08	non-syn	K	R
190854	116	G	A	0.92	0.08	syn	T	T
190977	157	C	A	0.9	0.1	non-syn	N	K
191520	338	C	T,	0.67	0.33	syn	D	D
191521	339	A	G	0.67	0.33	non-syn	N	D
192349	615	C	T	0.88	0.13	non-syn	H	Y
192385	627	G	A	0.88	0.13	non-syn	V	I
192768	754	T	A	0.89	0.11	non-syn	N	K
192771	755	A	G	0.89	0.11	syn	K	K
192798	764	A	T	0.89	0.11	non-syn	K	N
192801	765	G	A	0.9	0.1	syn	K	K
192836	777	A	G	0.86	0.14	non-syn	D	G
192871	789	A	G	0.86	0.14	non-syn	N	D
192885	793	T	C	0.86	0.14	syn	D	D
192889	795	G	A	0.86	0.14	non-syn	D	N
192906	800	T	C	0.78	0.22	syn	D	D
193218	904	C	T	0.91	0.09	syn	S	S
193224	906	A	T	0.9	0.1	non-syn	K	N
193480	992	A	T	0.83	0.17	non-syn	I	F
193483	993	A	G	0.83	0.17	non-syn	T	A
193537	1011	G	A	0.71	0.29	non-syn	E	K
193905	1133	C	G	0.67	0.33	non-syn	S	R
194355	1283	C	T	0.83	0.17	syn	F	F
194646	1380	T	A	0.9	0.1	non-syn	N	K
194675	1390	C	G	0.9	0.1	non-syn	S	C
195060	1518	C	T	0.9	0.1	syn	N	N
195090	1528	A	C	0.89	0.11	non-syn	E	D
195635	1710	A	G	0.9	0.1	non-syn	N	S
196248	1914	C	A	0.67	0.33	non-syn	N	K
196249	1915	A	G	0.6	0.4	non-syn	K	E
196269	1921	A	C	0.9	0.1	non-syn	K	N
196270	1922	G	A	0.9	0.1	non-syn	E	K
196332	1942	C	A	0.93	0.07	non-syn	N	K
196347	1947	T	C	0.92	0.08	syn	D	D
197219	2238	A	G	0.83	0.17	non-syn	K	R
197503	2333	T	C	0.73	0.27	non-syn	F	L
197583	null	G	A	0.71	0.29	non-coding	null	null
197594	null	T	A	0.9	0.1	non-coding	null	null
197596	null	A	T	0.73	0.27	non-coding	null	null
197604	null	A	T	0.57	0.43	non-coding	null	null
197786	null	A	C	0.9	0.1	non-coding	null	null
198697	2653	T	C	0.92	0.08	non-syn	V	A
198711	2658	G	A	0.9	0.1	non-syn	E	K
199475	2912	G	T	0.91	0.09	non-syn	K	N
199979	3080	C	T	0.82	0.18	syn	G	G
199987	3083	A	G	0.92	0.08	non-syn	H	R
200006	3089	T	C	0.7	0.3	syn	G	G
200012	3091	T	C	0.91	0.09	syn	N	N
200014	3092	G	A	0.73	0.27	non-syn	R	H
200020	3094	A	G	0.91	0.09	non-syn	N	S
200039	3100	C	T	0.56	0.44	syn	N	N
200047	3103	G	A	0.58	0.42	non-syn	S	N
200074	3112	G	A	0.67	0.33	non-syn	S	N
200376	3213	G	A	0.93	0.07	non-syn	V	I
201246	3426	T	G	0.78	0.22	syn	G	G

Table 5.15 Polymorphisms in PFA0220w gene in *P.falciparum* population

SNPs obtained by combining the data Broad Institute, NIH, and Wellcome Trust Sanger Institute (Jeffares et al., 2007; Mu et al., 2007; Volkman et al., 2007)

```

P.falciparum_PFA0220w      -----TNSELDYLFLLEIKSFFKNMLTTDKSYISADRVLNMLPVELNNR 3270
P.vivax_PVX_081540        TTPTNQVATAESERFLQELKNLYIIMTTTTKKTYVSPDCILGLLPHELNNR 510
P.knowlesi_PKH_020900     IPPMNHIGTVESERFLQELKNLYIIMTTTNKMYVSPDSILGLLPYELNNR 3342
P.chabaudi_PCHAS_020720   MNNNNNNANLLSKRFLYELKILFKLMTTTNKKYVSPDNILGILPQELNNR 2692
P.berghei_PBANKA_020880   PNINDSYTNLLSKRFLYELKILFKLMSTTNKKYVSPYSILSILPQELNNR 2716
P.yoelii_PY01242          --INDSHKNLLSKRFLYELKILFKLMSKTNKKYISPDNINILPQELNNK 1781
                          . . . * * * : : * . * * * : * : * * * :
                          . . . * * * : : * . * * * : * : * * * :

P.falciparum_PFA0220w      NQQDTEVFRYIFDKLGGSEKEFLRLIFSGVVIQKQCQKCLFISKKEEI 3320
P.vivax_PVX_081540        NQQDTELLRYIFDKLGGSSREFLRLIFSGVLVQKQCQKCFVSKKEEV 560
P.knowlesi_PKH_020900     NQQDTELLRYIFDKLGGSSREFLRLIFSGVLVQKQCQKCFVSKKEEV 3392
P.chabaudi_PCHAS_020720   NQQDTELFRTFEQLGGSEKKFLRLIFSGVVIQKQCQKCFVSKKEEI 2742
P.berghei_PBANKA_020880   NQQDTELFRTFEQLGGSEKKFLRLIFSGVVIQKQCQKCFVSKKEEI 2766
P.yoelii_PY01242          NQQDTELFRTFEQLGGSEKKFLRLIFSGVVIQKQCQKCFVSKKEEI 1831
                          *****: * * : * * : . . *****: : *****: : * * : * :

P.falciparum_PFA0220w      IHDLSFPVPDISTNEKLSIQKFFDTFIQKEKIYGNKKYKCSRCKNKRNALK 3370
P.vivax_PVX_081540        IHDLSFPVPKASERLSIQKFFDTFVQKEKIHGSKYRCSKCNKRNALK 610
P.knowlesi_PKH_020900     IHDLSFPIPVRASEKLSIQKYFDSFVQKEKIHGSKYRCSKCNKRNALK 3442
P.chabaudi_PCHAS_020720   IHDLSFHVPKASSKKQSIQKFFDTYIQKEKIYGNKKYKCSRCKNKRNALK 2792
P.berghei_PBANKA_020880   IHDLSFHVPKASTKKESIQKFFDTYIQKEKIYGNKKYKCSRCKNKRNALK 2816
P.yoelii_PY01242          IHDLSFHVPKASTKKESIQKFFDTYIQKEKIYGNKKYKCSRCKNKRNALK 1881
                          ***** : * : : : * * : * * : : * * : * * : * * : * * :

```

Figure 5.15 The fragment of the alignment of *ubp1* protein between six *Plasmodium* species. The position of V2697F and V2728F mutation are highlighted. The full alignment is available in Appendix D.

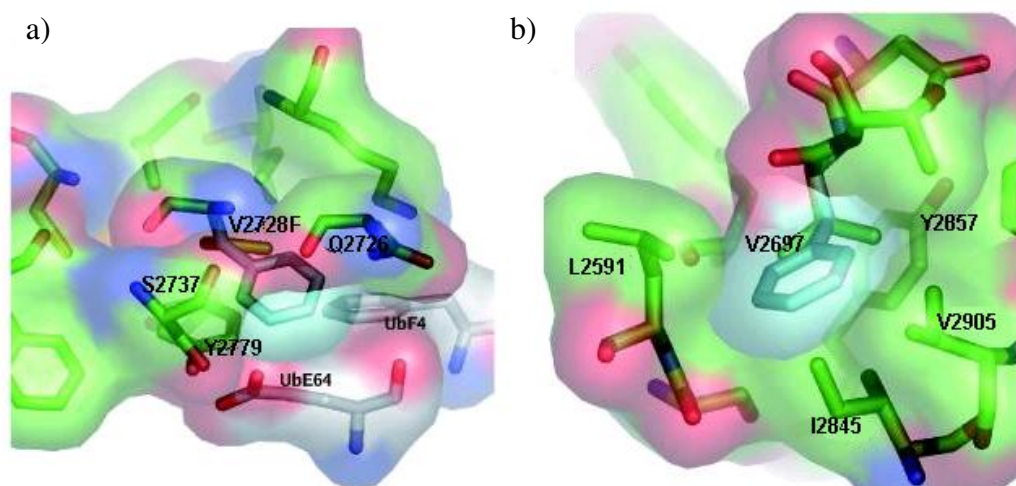


Figure 5.16 The predicted impact of the V2728F and V2697F mutations on the structure of the catalytic domain of *UBP1*

a) Residue environment around V2728F. The amino acid side chains are shown as sticks with carbons colored green for the *P. chabaudi* *UBP1* sequence and white for ubiquitin (UbF4, UbE64). The mutation to F2728, in its most likely orientation, is represented in stick form with carbons colored cyan. A surface representation is shown to indicate the tightness of fit around the phenylalanine.

b) Residue environment around V2697F. Side chains are shown as sticks with carbons colored green for the *P. chabaudi* *UBP1* sequence. The mutation to F2697 in the most likely orientation is represented in stick form with carbons colored cyan. A surface representation is shown to indicate the tightness of fit around the phenylalanine.

5.5. Discussion

5.5.1. The mutations within the resistant AS lineage

In this chapter Solexa genome resequencing was used to define the genetic changes that appeared in the AS lineage during the selection for chloroquine resistance. The good sequence quality allowed the reassembly of almost the whole genomes of both strains with the coverage high enough to discover the mutations. Moreover the bioinformatics analysis performed in parallel using two different types of software, identified the same mutation set which was later confirmed by dideoxy sequencing validation (while all changes classified as false positives were discarded). Therefore it is likely that the obtained inventory of mutations between the two strains is complete or nearly complete.

The main feature of this dataset is a very small number of mutations. Despite the multiple passages, mosquito transmissions and cloning events to which both strains were subjected, only 9 genetic changes (7 point mutations, 2 deletions) were identified between AS-sens and AS-30CQ. Moreover 6 out of the 7 point mutations are non-synonymous. Due to the compact nature of the *Plasmodium* genome we expected a high proportion of coding to non-coding mutations than in eukaryotic organisms with more complex genomes (Mu et al., 2002). However the bias obtained here suggests that many of the genetic changes in this lineage are not likely to be random events.

In fact at least five of them are connected with drug resistance phenotypes. The *dhfr* S106N was previously identified as determinant of pyrimethamine resistance in the lineage and my work associates *ubp1* V2728F, *aat1* A173E, *tm12* T719N and PCHAS_030200 T707N with CQ/ART resistance phenotype. Additionally also some of the other changes are situated within regions potentially connected to the pyrimethamine (chr 7), chloroquine (chr 5) and artemisinin (chr 7 and chr 14) resistance. Therefore the expected number of neutral mutations is very low if any.

5.5.2. The structure and acquisition of CQ and ART resistance

The combination of Solexa sequencing with the genetic data generated in chapter 4 defined the mutations contributing to the highest level of resistance to CQ – one mutation on chr 11 and one on chr 2, and two possible mutations on chr 3.

All of these were sequenced in AS lineage with an expectation that the order in which they appeared would reflect the order in which the chromosomes are selected with the increasing level of CQ resistance as presented in the model in Fig 4.17 (chapter 4). In fact, in line with these expectations, the A173E mutation in *aat1* was the first one to appear under CQ selection and is associated with acquisition of the first, lowest level of resistance. The acquisition of higher levels of CQ resistance however turned out to be more complex.

The genetic data suggest that chr 3 selection underlies an intermediate CQ resistance phenotype and that the chr 2 selection underlies the high level of CQ resistance. However, the analysis of the mutations present in AS-30CQ suggested otherwise, *i.e.* that the mutation on chr 2 in the intermediate AS-15CQ strain occurred before the mutations on chr 3 that were present only in highly resistant strain. The discovery of additional mutations in *tm12* and the *ubp1* gene and sequencing of additional strains in the lineage (AS-15MF and AS-ATN) introduced a further level of complexity, resulting in a complicated pattern of mutations on chromosome 3 and 2 presented in Table 5.11.

In summary, all strains characterized by ART resistance and higher level of CQ resistance (compared to AS-3CQ) contained one mutation in *aat1*, one mutation in *tm12* and one mutation in *ubp1*. However two alternative mutations in both *tm12* and *ubp1* were identified. These mutations are difficult to place in the lineage because they appear in two out of three clones selected from AS-15CQ. For each gene this means that either one of the strains is a revertant or that the same mutations appeared twice in the lineage in different clones. These explanations are considered highly unlikely.

These apparent contradictions may be resolved by acknowledging the fact that AS-15CQ is not clonal and as such cannot be assigned a reliable genotype. Therefore it can be a pool of various genotypes from which the three strains were

selected independently over time: AS-30CQ, AS-15MF and AS-ATN. In this context, it seems plausible that both mutations in *tm12* and in *ubp1* appeared during the selection for the intermediate CQ resistance in AS-15CQ. In fact only two genotypes present in the AS-15CQ could generate the observed genetic diversity (see Fig. 5.17). Two of the progeny strains (AS-ATN and AS-30CQ) would be generated by isolation of both original genotypes and the third (AS-15MF) would be a result of genetic recombination occurring whilst the AS-15CQ was transmitted through mosquitoes.

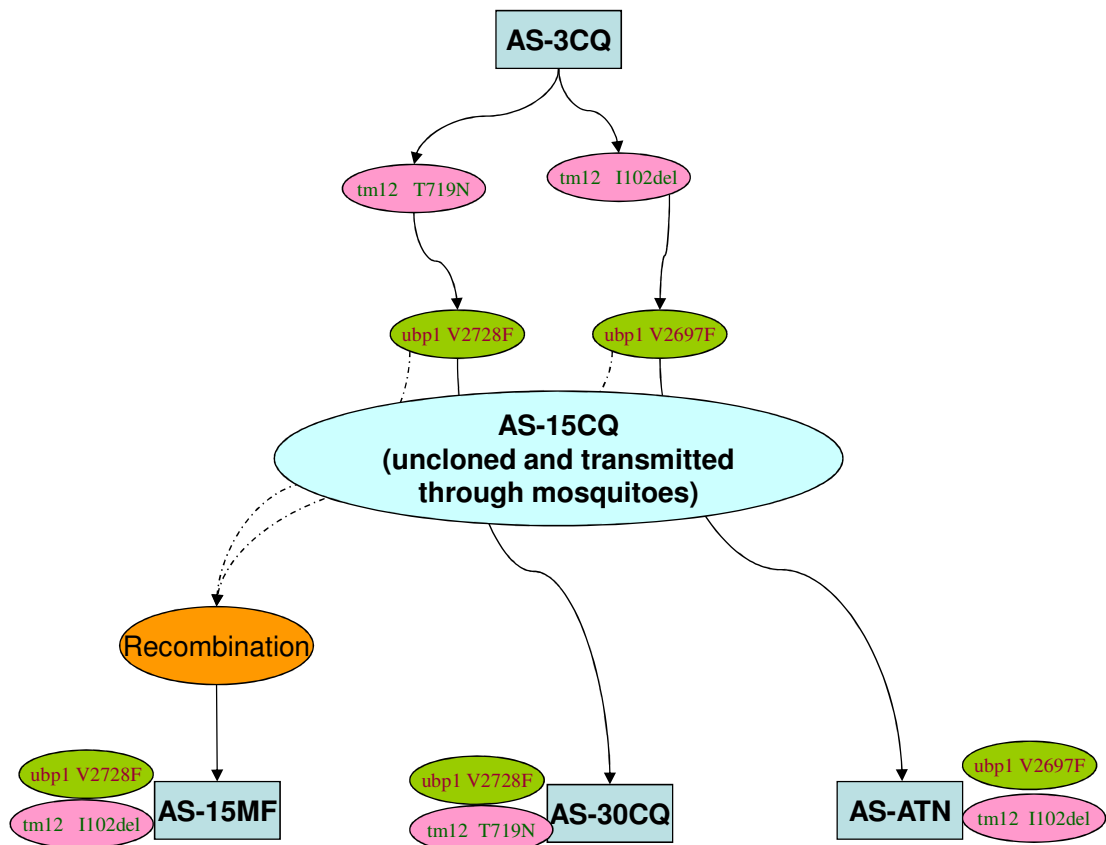


Figure 5.17 The possible order of occurrence of mutations on chr 3 and chr 2 in the AS lineage

This model resolves all the contradictions presented. It also allows for the possibility that the *tm12* mutations occurred before the *ubp1* mutations (in both cases), as predicted by the genomic data.

However it also leaves another unanswered question. If both *ubp1* and *tm12* contribute to the intermediate phenotype, what would be the genetic factor

responsible for further increase of CQ resistance between AS-15CQ and AS-30CQ? There are in fact two mutations specific for this strain: PCHAS_101550 Y162H on chr 10 and PCHAS_030200 T707N on chr 3. The first mutation failed to generate a selection signature in any of the experiments described in chapters 3 and 4 (none of the markers on chromosome 10 showed the signs of selection) and therefore its involvement in CQ resistance is unlikely. However PCHAS_030200 would be fit with obtained selection data. It would explain both the lack of additional selection valley after chr 2 (the selection of this gene it would be masked by the effect of *tm12* gene present on the same chromosome) and the lack of definitive selection valley on chr 3 (the presence of the two selected genes would cause the whole chromosome to be selected as observed in chapter 4 (Fig 4.12-14)). Therefore PCHAS_030200 is classified as a likely contributor to the resistance.

5.5.3. The possible mechanism of CQ/ART resistance in *P.chabaudi*

In this chapter I identified three (or possibly, four) novel genes potentially implicated in CQ resistance in *P.chabaudi*: *aat1*, *ubp1*, *tm12* and possibly *PCHAS_030200*. At least three of these genes (*aat1*, *ubp1*, *tm12*) contain coding mutations within highly conserved regions, affecting the residues that are conserved or semi-conserved in *Plasmodium* genus and likely to affect the function of the putative proteins.

Aat1, the main gene associated with the CQ resistance in this study, is a potential amino acid transporter localized into the digestive vacuole membrane. Interestingly, the two main determinants of CQ resistance in *P.falciparum* are the potential DV transmembrane transporters: *pfcr1* and *pfmdr1*. Additionally one of the possible functions of *PFCRT* is amino acid transport (Martin et al., 2005). Also the A173E mutation in *AAT1* bears some similarities to the crucial K76T mutation in *PFCRT*. In both cases, these mutations are located at the entrance to the predicted transporter. In the case of *aat1*, A173E is predicted to lie before the beginning of the 1st transmembrane domain on the digestive vacuole side. In the case of *pfcr1*, K76T lies at the beginning of the first TMD. Additionally, in both cases, the basic or neutral residue is replaced by a neutral or acidic one, respectively. This change

(increasing negative charge) can possibly act to facilitate the transport of the positively charged CQ (a weak base). It is therefore tempting to speculate that *AAT1*, similarly to *PFCRT*, either actively transports or allows the leak of CQ from the digestive vacuole.

The role *TM12* is more speculative. This protein acts epistatically to *att1* and is located in the same cellular compartment (DV). It is therefore possible that it participates in the same mechanisms, modifying the CQ transport. It may, for example, act as another transporter (similarly to *PFMDR1* in *P.falciparum*) or affect the function of *AAT1* directly (by interacting with that protein) or indirectly (alternating the DV environment, compensating for the loss of function caused by A173E mutation etc.).

The discovery of a third conserved gene (*ubp1*) connected with CQ resistance is of particular interest. It contributes to high-level CQ resistance without being directly involved in the transport and, most importantly, it also confers the ART resistance to the parasites, without the previous exposure to this drug. The fact that the same gene has an effect on two drugs of supposing different mode of action raises several questions. Ubiquitination and de-ubiquitination may regulate the post-translational state of many protein targets in many different ways *e.g.* affecting their function, targeting them to particular intracellular compartment or to degradation etc. (d'Azzo et al., 2005). This plays a major role in cellular processes such as cell-cycle control, protein turnover, intracellular trafficking of membrane proteins and receptor-mediated endocytosis. Thus a number of intriguing possibilities present themselves. Does *ubp1* change the ubiquitin state of any of the putative transporters connected with CQ resistance (*AAT1*? *TM12*?, *PFCRT* and *PFMDR1* homologues?)? Does it regulate their function? Or does it take part in general stress responses, allowing the parasite to survive drug-induced oxidative stress more successfully, by globally regulating protein turnover? Does it modify the cell cycle of the parasite allowing the temporary arrest in response to drug pressure (one of the main targets of HAUSP protein, a structural homologue of *ubp1* in human, is p53, a well-known protein involved in cell cycle control and apoptosis (Sarkari et al., 2010)? The identification of *ubp1* mutations therefore opens up new directions to investigate antimalarial drugs and the molecular mechanisms of their action and resistance.

6. The fitness cost of high CQ resistance.

6.1. Introduction

Section 1.3.3 of Chapter 1 describes the concept of fitness cost of drug resistance and reviews the evidence indicating its existence in malaria parasites. For example, the *pfcr*-dependent chloroquine resistance in *P.falciparum* is thought to be a significant burden to the parasite. However the study of this phenomenon in field isolates is often difficult due to ethical issues and the large number of additional factors affecting the performance of resistant and sensitive strains.

The *P.chabaudi* lineage presented here offers a unique opportunity to study the fitness consequence of CQ resistance *in vivo* under laboratory conditions, because the behaviour (growth) of a resistant mutant and its isogenic sensitive progenitor can be studied in parallel. In this study, it allows us to evaluate the effects of the mutations associated with a high level of resistance without the confounding effects of additional genetic variation.

The first indication of fitness costs related to high CQ resistance were observed in the AS-30CQ v. AJ and AS-3CQ v. AJ competition experiments described in Chapter 2. The highly resistant AS-30CQ strain grew slower than its low-resistant progenitor AS-CQ in both single and mixed infections. However a clear interpretation of the results was not possible because the parasites were tested in separate experiments, a few months apart, using different batches of animals, and slightly different treatment protocols. Moreover, in single infection controls only one mouse was used for each strain/treatment variation. Therefore, it was impossible to distinguish between differences in the parasites themselves, and host variation. Finally, in those mixed infections, the parasites were competed with a genetically different strain, AJ. The relative fitness of the two AS strains in such a situation may depend on many complex factors and interactions.

Therefore in this chapter competition experiments directly comparing the two congenic strains was performed in order to investigate whether a measurable difference of fitness exists between AS-30CQ and AS-3CQ during the mouse infections.

6.2. Experimental design

Note: This experiment, competing the high CQ resistant clone AS-30CQ with the low CQ resistant clone AS-3CQ, was attempted before there was a full inventory of mutations arising in the AS lineage. It was not possible to analyse the competition between AS-3CQ and AS-PYR1 because no discriminating genetic marker was available. However, one mutation, ubp1 V2728F discriminating AS-3CQ and AS-30CQ was known (Hunt et al. 2007).

The general experimental design is presented on Fig.6.1. The mixture of two parasite strains – the highly resistant AS-30 CQ and low resistant AS-3CQ was grown in 4 groups of 4 mice treated for the first three days (d0-3) with following doses of chloroquine:

- Untreated - where growth of both strain can be monitored without the drug pressure
- 1.5 mg of CQ kg⁻¹ day⁻¹ - the very low subcurative drug treatment that would provide the moderate level of stress for both resistant strains,
- 3 mg of CQ kg⁻¹ day⁻¹ - the concentration that allows the growth of both strains, but proved to be restrictive for the sensitive strain, AS-sens
- 10 mg of CQ kg⁻¹ day⁻¹ - the concentration that should differentiate between the high-resistant strain, AS-30CQ and the lower-resistant strain AS-3CQ.

In addition, in order to monitor the effect of the various drug doses, single infections were set up. AS- sens, AS-3CQ and AS-30CQ were grown in the presence of 0, 3 or 10 mg of CQ kg⁻¹ day⁻¹. Each treatment group contained 3 animals. All infections and parasitaemia were monitored by daily blood smears.

In order to monitor the proportion of the two strains in mixed infections, every second day a blood sample (tail bleed) was taken for quantitative genotype analysis. Mice 1 and 3 from each group were sampled on odd days (day 1, 3, 5 p.i etc.) and mice 2 and 4 on even (day 2, 4, 6 p.i. etc.) days. The proportional sequencing of part of the *ubp1* gene containing the previously identified mutation was used to estimate the proportion of resistant and sensitive parasites in each sample.

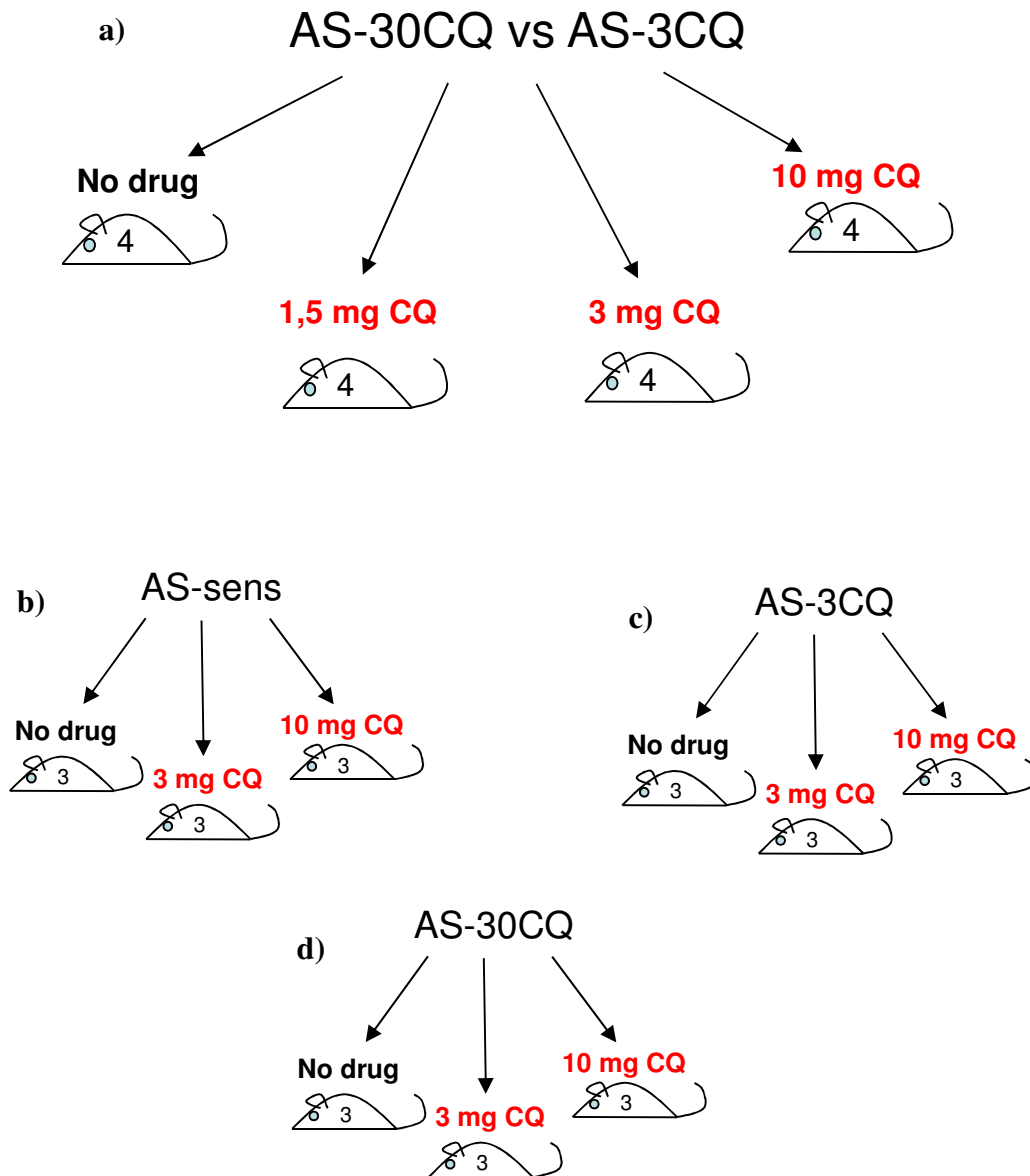


Figure 6.1 The experimental design of AS-3CQ v. AS-30CQ competition experiment

Each mouse represents a group of animals injected with the mixture of AS-3CQ and AS-30CQ strains (a), or the single AS-sens (b), AS-3CQ (c) or AS-30CQ (d) strain. The number inside the mouse silhouette represents the number of animals in the group

6.3. Materials and methods

6.3.1. Experimental infections

The AS-sens, AS-3 CQ and AS-30CQ strains were described in section 1.7.2 of Chapter 1. The inoculations, drugging, parasitaemia calculation, blood sampling and DNA isolation was performed as described in Chapter 2 Sections 2.3.2-2.3.5.

6.3.2. Proportional sequencing of ubp1 gene

The isolated DNA was amplified using a nested PCR. The inner and outer primers were as follows:

Outer:

Forward: 5'-ATGCAAACCTTACTTTCAAAACG-3'

Reverse: 5'-TTGTTGCATTTTCGAGCATTTG-3'

Inner:

Forward: 5'-CAAATAAAAAATATGTTTCACCAG-3'

Reverse: 5'-CGAGCATTTGTATTTATTGTTTCC-3'

Amplification was conducted as in Chapter 2, using an identical program to that used to amplify *dhps*. The products were purified using QIAquick PCR purification kit (Qiagen) and sequenced (using forward and reverse) using an ABI Prism0® BigDye™ Terminator dideoxy sequencing system (Applied Biosystems). All amplification reactions were performed in triplicate.

The resulting electropherograms were analysed using Chromas 2.33 software (Technelysium Pty Ltd) and the heights of polymorphic peak on the position 237/176 (forward/reverse chromatogram) was calculated as described in Chapter 2. The assay was calibrated using parasite mixtures of known proportion (data not shown).

6.3.3. Statistical analysis

The influence of various factors on parasitaemia and parasite (allele) proportion was tested with general linear model using R 2.8.1 software (<http://www.r-project.org/>) with lm4 package. The parasitaemia or parasite proportions were used as response variables. The combined results from all of the time-courses of infection were used with mouse number as a grouping factor. All data was arcsin-transformed to meet the criteria of normal distribution. The explanatory variables included 1) drug treatment – to compare the behaviour of the same strain at various drug dosed, 2) parasite strain – to compare the performance of different strains at similar drug concentration and 3) type of infection – to compare the behaviour of strain between the single and mixed infections.

6.4. Results

6.4.1. Single clone infections

The three strains, the sensitive AS-sens, the low resistant AS-3CQ and the highly resistant AS-30CQ were grown separately at 0, 3 and 10 mg of CQ kg⁻¹ day⁻¹. The parasitaemia of all infections were monitored and presented (Fig 6.2).

In *untreated animals* all three strains presented the growth pattern typical of rodent malaria, reaching peak parasitaemia after about 8-9 days followed by rapid clearance (Fig 6.2a). The sensitive strain AS-sens reached a higher peak parasitaemia ($51.6 \pm 4.6\%$) than the low resistant AS-3CQ ($36.01 \pm 1.88\%$) and the difference between the growth of two strains was statistically significant ($P = 0.037$). However no difference was observed between the growth of the low resistant and highly resistant clone (peak of parasitaemia $38.6 \pm 7.44\%$; $P = 0.27$)).

At **3 mg of CQ kg⁻¹ day⁻¹** (Fig 6.2b) the growth of sensitive AS-sens was significantly suppressed – the parasites appeared only at day 12 and did not reach a peak of infection during the course of experiment. The low resistant parasite appeared (day 7) and reached the peak of parasitaemia ($15.8 \pm 4.8\%$) on day 11. Interestingly the AS-30CQ parasite did not appear as early as AS-3CQ, appearing only on day 10 and reaching the peak of parasitaemia on day 13 with ($21.1 \pm 2.1\%$). The difference between AS-3CQ and AS-30CQ however was not statistically significant.

The 10 mg of CQ kg⁻¹ day⁻¹ treatment resulted as expected in complete suppression of growth for both the sensitive and low resistant AS-3CQ parasites. The highly resistant AS-30CQ managed to develop parasitaemia whose peak appeared on the same day (day12) and was only minimally decreased ($13 \pm 3.5\%$) in comparison with the 3 mg of CQ kg⁻¹ day⁻¹ infections.

In summary no significant growth difference could be observed between AS-3CQ and AS-30CQ except at the highest concentration of CQ.

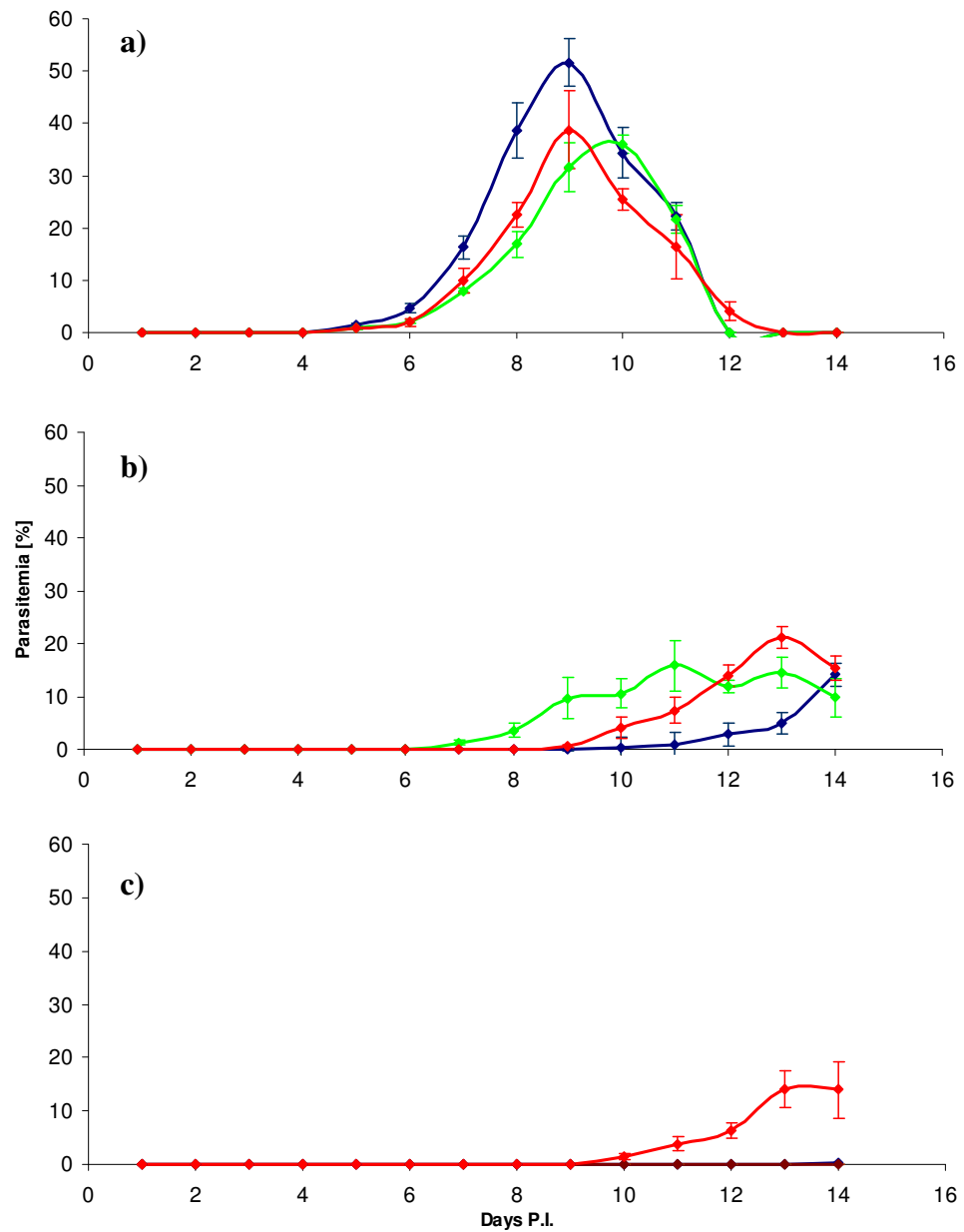


Figure 6.2 The single strain infections

The growth of AS-sens (◆), AS-3CQ (◆) and AS-30CQ(◆) in untreated infections (a) and at 3 (b) or 10(c) mg of CQ kg⁻¹ day⁻¹. The average results from three animals are shown.

6.4.2. Mixed infections

In order to directly compare the low resistant and highly resistant strain, mixed infections containing equal proportions of AS-3CQ and AS-30CQ were established and treated with 0, 1.5, 3 and 10 mg of CQ kg⁻¹ day⁻¹. Their parasitaemia and the proportions of the strains in the infection were monitored and are presented in Fig 6.3

6.4.2.1. Parasitaemia

All animals developed infections. The growth of the mixtures in the untreated and 1.5 mg of CQ kg⁻¹ day⁻¹ animals (Fig.6.3 a-b) was comparable to the growth of both clones in untreated single infections – the parasitaemia appeared on day 4 and reached ~ 35% on day 8 before a gradual decrease. No significant difference was observed between the growth in 0 and 1.5 mg of CQ kg⁻¹ day⁻¹ groups.

In contrast, as expected from previous experiments, the higher treatments (3 and 10 mg of CQ kg⁻¹ day⁻¹) (Fig.6.3c-d) significantly decreased parasite growth ($P=0.037$ and $P= 0.001$ respectively). In those groups the peaks of parasitaemia appeared later (days 12 and 14 respectively) and were lower ($19.4 \pm 3.4\%$ and $14.2 \pm 2.7\%$) than in either untreated population.

6.4.2.2. Proportion of AS-3CQ and AS-30CQ

An analysis of the blood sample used to inoculate mice with mixed infections showed that the percentage of the highly resistant strain AS-30CQ was 49.77 ± 0.2 in line with expectations.

Thereafter in all experimental groups, the percentage of AS-30CQ deviated depending upon the dose of drug used (Fig 6.4). The percentage of AS-30CQ dropped in the groups treated with between 0 and 3 mg of CQ kg⁻¹ day⁻¹ but rose in the group treated with 10 mg of CQ kg⁻¹ day⁻¹ during the first 3 days of infection. Interestingly however, in contrast to previous experiments using competition between pyrimethamine resistant and sensitive parasites, which reported the constant decrease of the proportion of resistant parasite throughout the infection (Walliker et al. 2005), there was little further change during the rest of the infection.

Specifically, consistent with previous phenotyping data, at the highest CQ

concentration - 10 mg of CQ kg⁻¹ day⁻¹ – after day 3 p.i. the infections were dominated by the more resistant strain (Fig. 6.3d). In all animals, 100% AS-30CQ was observed during the course of the infection, except at days 5 to 7 when some 5-20% of AS-3CQ was observed.

At 3 mg of CQ kg⁻¹ day⁻¹ (Fig. 6.3c and 6.4) the proportion of the two strains was more equal, resulting in ~ 30% of AS-30CQ throughout the infection. The variation between individual animals was greater here than in other treatment groups. For instance the average proportion of AJ alleles throughout the infection in mouse 1 showed about 7.2% ± 0.02 of AS-30CQ while in mouse 3 it was 45.8% ± 0.6 (Fig 6.4). Within each animal, though, the proportion was still remarkably consistent, suggesting that the proportional sequencing technique was reliable and suggesting that there was consistent variation in clone proportion between individual mice.

In untreated animals (Fig 6.3a) the proportion of the highly resistant AS-30CQ strain in the infection remained at ~10 – 20% from day 3 p.i. onward. This shows that in a drug free environment AS-30CQ performs worse than AS-3CQ even if no significant difference of growth between the clones could be observed in single infections. This is suggestive of a ‘fitness cost’ of the mutations identified in Chapter 5

Interestingly, the low sub-curative concentration of chloroquine - 1.5 mg of CQ kg⁻¹ day⁻¹ there was a stronger shift of proportion in favor of the low-resistant clone – on average only 6.4 ± 0.9% of infection consisted of the highly resistant AS-30CQ (Fig 6.3b). This value was significantly (P=0.018) lower than in the untreated group, suggesting that the difference of growth between the two strains is stronger at low sub-curative concentration than for the untreated controls.

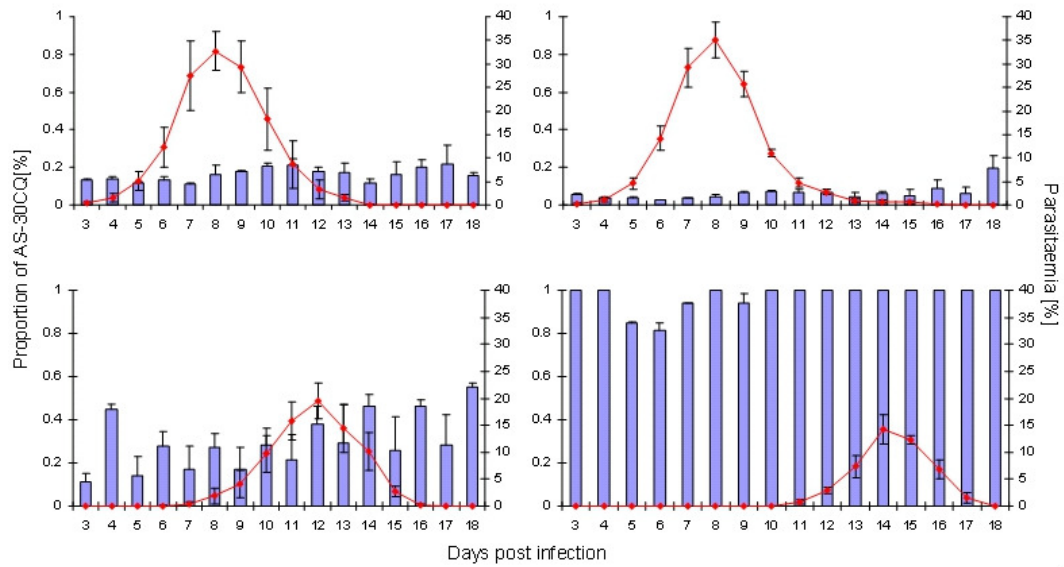


Figure 6.3 AS-30CQ +AS-3CQ mixed infection

The parasitaemia (red lines) and proportions of AS-30CQ strain (blue bars) in mixed AS-30CQ v. AS-3CQ infection treated with 0 (a), 1.5 (b), 3(c) and 10(d) mg of CQ kg⁻¹ day⁻¹ of CQ. Each graph is an average of four animals.

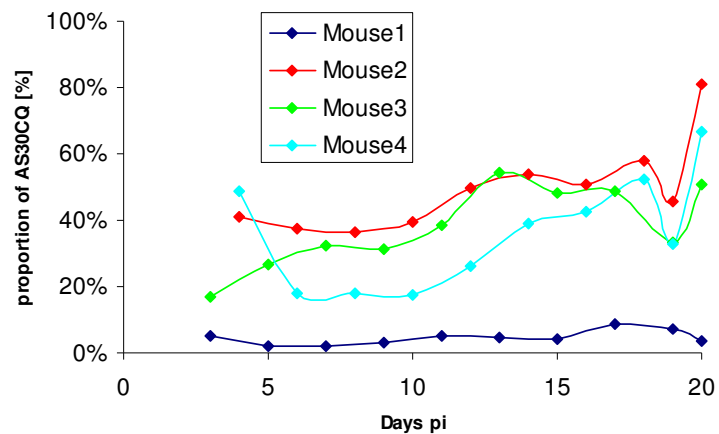


Figure 6.4 The variation with animals treated with 3mg of CQ kg⁻¹ day⁻¹

The proportional sequencing of 3 mg of CQ kg⁻¹ day⁻¹ treated mice. Each mouse is represented by different color.

6.4.2.3. Partial infections

For each group of infections, we were able to calculate a clone-specific ‘virtual’ parasitaemia by multiplying the total parasitaemia by the proportion of that clone. It allowed us to compare directly the growth of a strain in mixed and single infection at the same drug concentrations (Fig 6.5)

In untreated mixed infections, the growth of AS-3CQ is very slightly decreased relative to its growth in single infections (difference not significant) while AS-30CQ is significantly reduced in the presence AS-3CQ, achieving only around 1/6 of the parasitaemia obtained in single infections. This effect was even stronger in the 1.5 mg of CQ kg⁻¹ day⁻¹ and 3 mg of CQ kg⁻¹ day⁻¹ groups (the 1.5 mg of CQ kg⁻¹ day⁻¹ data were compared to *untreated* single infections because single infection data at 1.5 mg of CQ kg⁻¹ day⁻¹ was not available). The low resistant AS-3CQ strain showed no differences in growth between mixed and single infections at these concentrations whereas the highly resistant AS-30CQ was significantly decreased (at 3 mg of CQ kg⁻¹ day⁻¹) or almost completely suppressed (at 1.5 mg of CQ kg⁻¹ day⁻¹) compared to its behaviour in single infection. Only at the highest dose of treatment (10 mg of CQ kg⁻¹ day⁻¹) that was eliminating its competitor, AS-30CQ, was the parasitaemia comparable to those in single infection. Even there however, the presence of competitor delayed the peak of parasitaemia for two days.

In summary, the low-resistant strain remains unaffected by the presence of competitor and the changes in proportion of parasites are due to various degrees of suppression of highly resistant AS-30CQ strain.

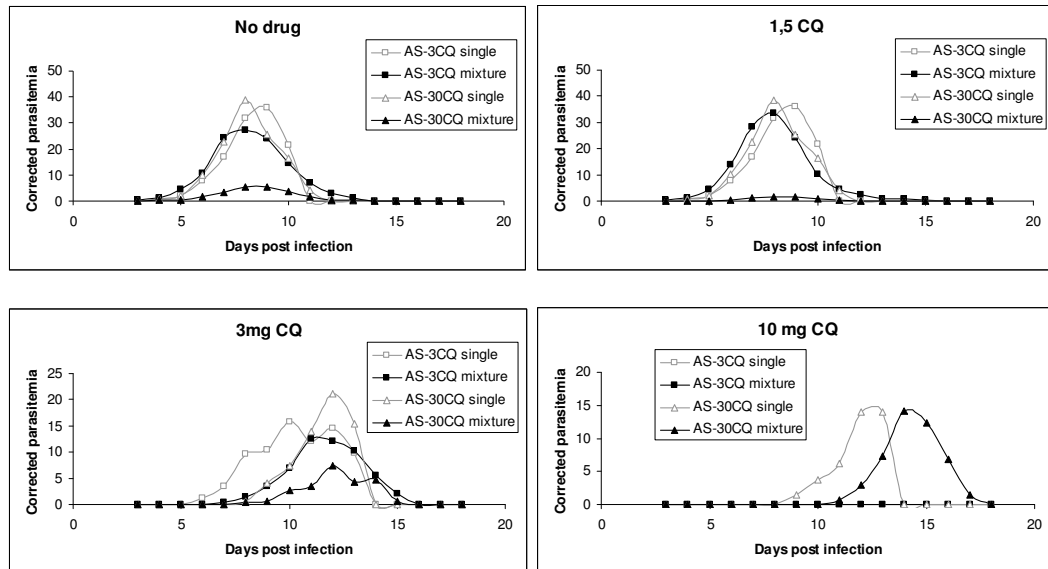


Figure 6.5 Differences between single and mixed infections

The comparison of behaviour of AS-30CQ and AS-3CQ strains in single (grey lines) and mixed (dark lines) infections under different concentrations.

6.5. Discussion

6.5.1. Do mutations underlying chloroquine resistance incur fitness costs?

In this study I compared the growth in mixed infections of two isogenic clones of the malaria parasite, *P. chabaudi* in a rodent host, namely the low resistant AS-3CQ and the highly resistant AS-30CQ. As expected, AS-30CQ outgrows AS-3CQ at high CQ doses confirming the increase in CQ resistance in AS-30CQ relative to AS-3CQ. In contrast, in the absence of drugs, it is the low resistant AS-3CQ that grows significantly better than highly resistant AS-30CQ. At low, sub curative concentrations the AS-3CQ strain outgrows AS-30CQ even more than in the untreated infections; almost completely suppressing its competitor. At an intermediate concentration, we observed a more equal growth of both strains although the low resistant one still achieves the higher parasitaemia.

These data can be interpreted as follows. The low-resistant clone AS-3CQ has increased ‘host-specific’ fitness relative to the highly resistant AS-30CQ at low and intermediate doses of chloroquine as well as in the absence of drug. Only at the high doses does the advantage of the high CQ resistance outweigh the decreased ability to compete. At intermediate concentration the two effects balance each other, possibly causing the final result to be more susceptible to random changes, slight variation in CQ availability within the host, or experimental error (hence the higher variation between the animals in this group). Therefore these results taken together strongly suggest the existence of fitness cost associated with mutations conferring higher levels of CQ resistance in rodent malaria.

Here it has to be noted however that these data concerns only the “rodent part” of the malaria life-cycle. In order to fully assess the impact of resistance on reproductive success the other parts of the cycle including gametocyte production, mosquito transmission, sporozoite replication and mammalian transmission etc. must be taken into account. Therefore further research is necessary in order to understand fully the impacts of high level of resistance on parasite fitness. Nevertheless, it is

possible to see how a similar analysis is possible after parasite mixtures have been passed through both mammalian and mosquito hosts. Indeed, because genetic recombination will occur between the high- and low-resistance strains, independent estimations of the changes in the proportions of all specific mutations is possible.

6.5.2. What genetic changes account for the fitness costs?

If a fitness cost is a stable feature of the highly resistant strain, it is conferred by the mutations that took place between AS-30CQ and its sensitive progenitor AS-3CQ. In Chapter 5 I described the use of Solexa whole-genome re-sequencing to define a complete inventory of SNPs arising in the AS drug-resistant lineage (Table 5. 5 in that Chapter and Appendix D).

In summary there are four point mutations which arise between AS-3CQ and AS-30CQ. Three of them (point mutations in *ubp1* gene on chr 2, *tm12* on chr 3 and non-syntenic PCHAS_030200 on chr 3) are non-synonymous mutations linked to high-level chloroquine resistance. The others are a non-synonymous SNP in PCHAS_101550 gene on chr 10 and a large deletion including PCHAS_051920 gene on chr 5. The roles (if any) of these two mutations in drug responses cannot be defined at present.

One interpretation is that one or more of the mutations connected with resistance incur fitness costs and that the two remaining mutations are either neutral or partially compensate those costs. Therefore a fitness cost observed in AS-30CQ would be a direct consequence and inherent feature of resistance. Alternatively, the fitness costs may be incurred by one of unselected mutations. In that case it would be random consequence of bottleneck that appeared in malaria population during the selection of resistant genotype (“a founder effect”). The second hypothesis appears less likely as such a mutation (conferring a significant fitness cost without any advantages related to drug resistance) would be very likely to be lost during the passages. However further research will be required to confidently assign the fitness cost of specific mutations. For example the generation of recombinant parasites carrying different combinations of the five described mutations and a series of competition experiments between them could address this question. Alternatively, parasites AS-3CQ and AS-30CQ could be cycled through mice and mosquitoes.

Recombination would allow ‘shuffling’ of the mutations and hence their fate in the absence of drugs could be monitored independently of each other.

6.5.3. Additional observations

6.5.3.1. The advantage of low resistant strain at low drug concentration.

Because no treatment privileged the growth of the low resistant strain and the high treatment dose resulted in infection composed almost exclusively from the highly resistant one, I was expecting that the intermediate treatment doses would result in a gradual transition between these two states. Instead, the proportion of the more resistant strain reached a minimum value in animals treated with 1.5 mg of CQ $\text{kg}^{-1} \text{ day}^{-1}$. This suggested that the difference of fitness is stronger at low, sub-curative concentration than in untreated animals.

One of possible explanations of this phenomenon is that the low drug dose does not activate the resistance mechanism but induces a low level stress in both strains, so enhancing the competition between the two clones and putting the more “fragile” clone at greater disadvantage. This would be consistent with recently published findings suggesting that between two sensitive clones the less virulent one is going to be less affected by drug (Schneider et al., 2008).

6.5.3.2. The stability of parasite proportions over time.

In previous competition experiments, including pyrimethamine-resistant v. pyrimethamine-sensitive clones described in (Walliker et al., 2005) and AS-30CQ v. AJ and AS-3CQ v. AJ experiments described in Chapter 2, the proportion between the two parasite strains changed steadily throughout the infection. It suggests that the fitness cost is maintained constantly during the exponential growth of parasite and as the infection is controlled by host factors such as immunity, red blood cell availability etc.

In contrast, during the current experiment, the proportion of both clones changed rapidly during the first three days before stabilising later in the infection, during days 4 - 18. It implies the existence of some mechanism that maintains an equilibrium value, such as the existence of two opposing but balanced processes. Alternatively, it is possible that for first three days only, the strains are in exponential

growth phase and after that their growth is limited by the available resources (e.g. erythrocytes) rather than their intrinsic growth rate which may be genetically determined. This might suggest that fitness costs only prevail while the growth rate is high, the parasitaemia is low, a cytokine storm is small or the immune response negligible etc. Thereafter, it is possible that there are no differences in growth rates between the clones.

In summary the generated data strongly suggest the existence of a fitness cost related to the high CQ resistance in AS lineage and reveals some previously unreported phenomena associated with mixed infection. These processes can be further investigated, leading to an improved understanding of parasite ecology and possibly better treatment policies.

7. General discussion and final remarks

In this Chapter I wish to summarize the most important findings of this thesis and their relevance for malaria research as well as indicate some future research directions suggested by the findings from this study.

7.1. Summary of results

The main objective of this thesis was to identify the genetic basis of CQ and ART resistance in the AS lineage of rodent malaria parasite *P.chabaudi*. The selection conditions allowing me to discriminate between the sensitive, low resistant and highly resistant parasites of the lineage were established in preliminary experiments (Chapter 2) and used for the drug-selection of genetic cross and backcrosses between the highly resistant (AS-30CQ) and sensitive (AJ) parasites (Chapter 3 and 4). The genetic analysis of this material generated genome-wide scans of drug-selection, mapping three critical loci contributing one after another to increasing level of CQ resistance (on chr 11, chr 3 and chr 2) and one major locus on chr 2 responsible for ART resistance.

Next Solexa genome re-sequencing was applied to identify a comprehensive inventory of mutations arising within the AS lineage (Chapter 5). The main result was that the pre-dominant selection valleys for different levels of CQ resistance and for ART resistance contained only one or two mutations. It was therefore possible to identify the A173E *aat1* and V2728F *ubp1* as the possible determinants of CQ- and ART-resistance, respectively. Additionally two genes on chr 3 (PCHAS_31370 and PCHAS_030200) and the *ubp1* mutation may contribute to higher levels of CQ resistance.

Finally the effect of the high CQ resistance on parasite fitness was studied in Chapter 6, suggesting that the highly resistant parasites are at a competitive disadvantage in the absence of drugs and at the low CQ concentrations.

7.2. Combination of LGS and Solexa sequencing for the identification of genes involved in drug resistance

The main scope of my work was to identify novel genes involved in drug resistance. The strategy used in this thesis exploits two novel approaches. Firstly, we adopted a QTL modification of LGS. This applied different levels of selection to achieve an elegant dissection of the genetic determinants of different levels of CQ resistance. Secondly, the massive power of new genomic sequencing technologies was applied to the drug-resistant AS lineage to generate a comprehensive inventory of mutations accumulating under selection by multiple drugs.

These approaches represent a substantial improvement over previously used strategies. Until this study, LGS was shown to be efficient in mapping only one dominant locus conferring the trait under investigation (Culleton et al., 2005; Martinelli et al., 2005; Hunt et al., 2007) and failed to supply the gene-candidates (unless a previously identified functionally relevant gene was present within the identified locus). The study of chloroquine resistance described here is the first case of successful use of LGS to identify multiple loci contributing to a given trait *and* the candidate mutations within them. Moreover, the use of different drug concentrations during LGS selection allowed not only the identification of all involved loci but also the evaluation of the strength of selection on each and the formulation of some hypotheses regarding the epistatic interaction between them; for example, the interaction between the chromosome 11 and 3 in case of higher level CQ resistance.

,

The chosen strategy proved to be a very efficient tool, allowing to pass from the characterisation of the phenotype (chapter 2 and 6), through the mapping of the involved loci (Chapters 3 and 4) to identification of candidate mutations (Chapter 5) within a relatively short amount of time. The whole process could be optimised and completed within less than 6 months. Additionally, once established, this strategy can be successfully used for identification of the genes involved in any phenotype that can be successfully selected in laboratory conditions. For example, this approach was also used concurrently to investigate the genetic basis of mefloquine and artesunate resistance. Although initially developed in malaria parasites, it can be

applied to other organisms provided that genetic recombination and selection of the desired phenotype is possible. In fact a very similar strategy was recently used in budding yeast to map the genes involved in different traits including resistance to ethanol, 4-nitroquinoline neomycin etc. (Ehrenreich et al., 2010).

7.3. Relevance for human malaria

7.3.1. The new gene candidates involved in CQ/ART resistance.

The first immediate consequence of the findings described in this thesis is the discovery of new gene candidates connected with CQ and ART resistance that may now be investigated in human parasites. Previous work has shown that the genetic determinants of drug resistance are often shared by *P.chabaudi* and *P.falciparum* (as in the case of pyrimethamine (Cowman and Lew, 1990), mefloquine (Cravo et al., 2003) and atovaquone (Afonso et al., 2010) resistance). Therefore it is possible that also the genes identified here will play some role in CQ/ART responses in human malaria.

In *P.falciparum* the main determinants of CQ responses were already identified (see section 1.4.2.) and no strong signatures of selection connected with CQ resistance could be detected on the chromosomes containing the homologues of *aat1*, *tm12* and *ubp1* genes (Mu et al., 2007; Patel et al., 2010). However, the identified genes may still be the additional factors modifying the CQ responses whose presence was described in section 1.4.2.3. In fact *aat1* homologue in *P.falciparum* was one of the 35 genes differentially regulated in three *pfcr*t mutants characterized by different chloroquine responses (Jiang et al., 2008) indicating its involvement in the CQ resistance pathway. More importantly however the discovered genes may be the main determinants of the CQ resistance in *P.vivax*, which is filogenetically closer to rodent parasites than *P.falciparum* (Silva et al., 2010) and which is able to acquire the CQ resistance in *pfcr*t-independent way (Nomura et al., 2001). The model of chloroquine resistance in *P.vivax* would have a great value, as despite the recent advances in culturing and genetic analysis, this parasite is still challenging to study in the laboratory conditions.

Additionally the involvement of *ubp1* (and possibly *aat1*) in the emerging ART resistance in *P.falciparum* remains to be investigated. Although this gene was not mutated in laboratory selected ART resistant *P.falciparum* (Chavchich et al., 2010) and the two mutations described in this study were not identified in field (Imwong et al., 2010) it is still possible it influences ART responses.

7.3.2. The insights into the mechanism of CQ and ART resistance

Despite the differences in the involved genes, parallels can be drawn between the putative mechanisms of CQ resistance in *P.falciparum* and *P.chabaudi*. The similarities between the main determinant of CQ resistance in *P.falciparum* (*pfcrt*) and *aat1* gene discovered here were already mentioned (see section 5.5.3). Both genes are predicted to be putative amino acid transporters situated within the digestive vacuole membrane. The mutation connected to CQ resistance is located on the internal side of digestive vacuole, outside the first transmembrane domain, and at the entrance to the putative channel. In both cases the CQ resistance seems to be connected with either removal of positive charge (*P.falciparum*) or introduction of the negative charge (*P.chabaudi*) that can facilitate the efflux of charged CQ from the vacuole. Additionally, in both systems a supplementary protein with 12 transmembrane domains, situated within digestive vacuole membrane, was identified – *pfmdr1* (another potential CQ transporter) in *P.falciparum* and *tm12* in *P.chabaudi*. In both cases this protein seems to act epistatically, increasing the level of resistance in parasites already carrying the crucial change within *aat1* or *pfcrt* gene. Finally, in both parasites CQ resistance is connected with CQ accumulation in resistant parasites (Miki et al., 1992), changes in the morphology of digestive vacuoles (Ohsawa et al., 1991) and reversal of resistance by calcium channel blockers like verapamil (Tanabe et al., 1990). Therefore, although different proteins may be involved, it is tempting to speculate that the same general resistance mechanism, based on CQ transport by modified amino acid transporter, exists in both species.

The involvement of *ubp1* in CQ/ART resistance doesn't have its direct equivalent in *P.falciparum*. Interestingly however, the study of ART resistance in *P.falciparum* reveals that it may be connected with delayed recrudescence and cell cycle arrest at the ring stage (Teuscher et al., 2010; Witkowski et al., 2010). As

mentioned previously (see section 5.5.3.), one of the possible functions of *ubp1* (based on the proteins it interacts with in yeast two-hybrid system) is cell-cycle control. Additionally, its structural homologue in human (HAUSP) is regulating the function of *P53* protein – a well-known mediator of cell cycle arrest in eukaryotic cells (Sarkari et al., 2010). It raises the question, whether also in the case of artemisinin the same general mechanisms of resistance may exist in both parasites.

In summary, regardless of their involvement in CQ/ART resistance in human malaria, the identity of the discovered genes supplies an interesting insight into the mechanisms of drug resistance in the *Plasmodium* genus that can be now further investigated in human parasites.

7.3.3. The connection between chloroquine and artemisinin resistance

In the light of the above similarities a connection between the CQ and ART resistance is of particular interests. Our findings strongly suggest that *ubp1* (and possibly *aat1*) is contributing to the resistance to both CQ and ART. Additionally in AS lineage, despite repeated efforts, artemisinin resistance could not be selected from the sensitive AS-sens strain. However it did appear spontaneously under CQ pressure after which this parasite could be further selected for increased ART resistance (see section 1.7.2). It suggests ART resistance (most likely conferred by the *ubp1* gene) can arise only on CQ resistant background.

It is not the first case of cross-resistance between those two drugs observed in rodent model. For example attempts of selecting ART resistant *P.berghei* (that was CQ-sensitive) failed. However the parallel experiment was successful in naturally CQ resistant *P.yoelii* resulting not only in appearance of ART resistance but also in spontaneous 7-fold increase of CQ resistance in that parasite (cross resistance with other quinoline have also been observed) (Chawira et al., 1986). Also the selection of *P.vinckei* for resistance to arteether (one of the artemisinin derivatives) resulted in spontaneous increase resistance to some of the quinoline drugs (CQ was not tested in this experiment) (Puri and Chandra, 2006).

In human malaria the relationship between the two drugs is less obvious. On one hand no connection can be observed between the main determinant of the CQ

resistance (*pfcr1*) and artemisinin responses. In fact, if anything, the contrary can be observed – an introduction of resistant *pfcr1* allele into sensitive parasites seems to be associated with increased sensitivity to artemisinin *in vitro* (Sidhu et al., 2002). and in field the ART treatment often selects for the sensitive *pfcr1* allele (Sisowath et al., 2009). On the other hand however, the secondary CQ-resistance determinant (*pfmdr1*) has been connected to changes in ART responses both *in vitro* and *in vivo* (see section 1.5.3.1.), although is unlikely that this gene is a sole determinant of ART responses. Additionally, the generation of artemisinin resistance in laboratory conditions does not seem to require the presence of CQ resistant background (Chavchich et al., 2010), however in field the resistance to artemisinin did arose in a region where the high CQ resistance reached nearly fixation (Lim et al., 2003; Saito-Nakano et al., 2008; Rungsihirunrat et al., 2009) and the recent study of the progeny of the cross between CQ resistant and CQ sensitive parasite (Beez et al., 2010) indicates that some genotypes will influence the parasite ability to develop stable ART resistance.

Are all these cases the correlation between quinolines and artemisinins due to a general “drug-response” phenotype mediated by *pfmdr1* and/or *ubp1*? Or some fundamental similarities between the two groups of drugs are still to be discovered? Is CQ resistance somehow priming the parasites to acquire the ART resistance? The answer to these questions would have very significant implications for international treatment policies as the resistance to chloroquine is still widely spread throughout the *P.falciparum* population and artemisinin is currently the first line of treatment.

7.3.4. A platform for identification of genes connected to drug resistance

In addition to the findings related with the two particular drugs chosen in this study, the approach developed in this thesis is a useful tool for feature identification of gene-candidates connected to the new drugs. As described previously the current strategies of gene identification in human malaria are ethically and technically difficult and often result in the identification of causative genes years after the resistance has been identified field. The rodent parasites, LGS and Solexa sequencing can be used for identifying candidate-genes involved in the resistance to new drugs as even before they are introduced. The identified candidates could be immediately

tested in human malaria (using transfection studies) and the mechanisms by which they are causing the resistance could be investigated.

The described approach could identify potential markers and mechanism of resistance before it actually arises in the field. This would allow the prospective monitoring of the population for signs of resistance and reacting before it reaches significant levels in the population.

7.4. Feature directions

This section briefly outlines the feature work based on the data from this thesis.

7.4.1. The use of Solexa SNPs quantification to improve the existing gene finding strategy.

The LGS ‘read-out’ can be improved by replacing pyrosequencing markers with quantitative SNPs defined and measured by Solexa sequencing. Previous work (Grech et al., 2002) suggests that may be ~100,000 SNPs differentiating the two parental parasite strains, AS and AJ. Therefore if both parents (AS and AJ) are sequenced, SNPs between the two parental strains can be defined and their frequency in drug-selected and unselected cross-progeny populations tested. Every SNP can be sequenced multiple times (20 - 60x) providing large amounts of LGS data which can be used to estimate errors in proportions and to define the limits of the base of the selection valley at greater resolution.

Update: at the time of submission of the final version of this thesis, such sequencing was preformed on the DNA from the backcross II treated with 3mg of CQ kg⁻¹ day⁻¹. As predicted ~100,000 SNPs were discovered and their frequency measured. The data analysis confirmed the results obtained with pyrosequencing markers and supplied better resolution of identified selection valleys. This data is being now subjected to further analysis.

7.4.2. Further investigation of gene candidates

The described work identified a number of genes potentially involved in CQ and ART response. Molecular biology approaches can now be used to unambiguously confirm the causative role of the described mutations in drug

resistance and elucidate the mechanisms by which they influence CQ/ART responses. Various *P.chabaudi* and/or *P.falciparum* parasites containing the mutated version of one or more of these genes can be generated using transfection technology and studied using a range of techniques involving growth inhibition tests, protein expression analysis, CQ/ART accumulation assays etc.

7.4.3. Exploration of fitness cost and its consequences

In order to explore the idea of fitness cost, a series of experiments can be designed during which mixtures of various parasites can be grown repeatedly through the full life-cycle of natural mice/mosquito passages and the parameters of the infection (strain proportion, anaemia, parasitaemia, number of oocysts etc.) recorded. This would enable us to estimate overall fitness costs (through the full life-cycle) and to establish where these costs have maximum effect. Alternatively genetic crosses can be used for the same purpose: in these cases the fitness implication of single mutations could be studied.

In summary, the described work offers multiple possibilities for further investigation. They offer exciting new options for malaria studies and can be of the utmost importance in understanding and limiting the evolution of drug resistance and of maximizing the impacts of malaria chemotherapy.

Appendix A – Buffers and Solutions

Citrate Saline

0.85 % w/v NaCl

1.5% w/v Tri-Sodium Citrate

Ph buffered to 7.2

PBS

0.8% w/v NaCl

0.02% w/v KCl

0.115% w/v Na₂HPO₄

0.02% w/v KH₂PO₄

Ph buffered to 7.4

Deep-freeze solution

:28%w/v glycerol,

3%w/v sorbitol,

0.65%w/v NaCl

sterilized by filtration

Mammalian Ringer Solution.

27mM KCl

0.15M NaCl

0.2mM CaCl₂

Ringer/FCS solution

50% heat inactivated foetal calf serum

50% mammalian Ringer Solution

20U/ml heparin

TBE buffer

100 mM TRIS

100 mM Boric Acid

2 mM EDTA

Parasite Lysis Buffer

150mM NaCl

25mM EDTA

0.25% SDS

0.125mg/mlProteinase K

Appendix B – General PCR protocol.

The following PCR protocol was using for DNA amplification :

The reaction was set in 100 µl tube as follws:

Ingredient	Concentration	Volume [µl]
DNA template	~40ng/µl	1
PCR Buffer (Promega)	5X	10
Mg Cl ²	25 mM	3
dNTP (Promega)	10 mM each	1
Primer 1 (forward)	10 pmol/µl	1
Primer 2 (reverse)	10 pmol/µl	1
Taq plymerase (Promega)	5U/µl	0.25
H ² O		32.75

Next the mixture was amplified using a standard PCR program (unless specified otherwise) :

Temperature	Time	
96°C	1 min	
96°C	1 min	x 30
52°C	1 min	
65°C	1 min	
65°C	10 min	
4°C	∞	

10 µl of generated product was used for electrophoresis in 1% agarose gel in TBE buffer with the addition of 1 µl od Safeview© stain. The gel was visualised using BioRad GelDoc 1000 UV transilluminator in order to confirm the presence of PCR product.

Appendix C – The list of pyrosequencing assays used in LGS analysis

The list of the assays used in LGS experiments described in Chapters 3 and 4.

The markers were designed as described in section xxx. For each selection, initially 2-4 markers on each chromosome were chosen in order to assess the overall shape of the genome scan. Later additional assays were added starting with the regions of particular interests (e.g. potential selection valleys). The addition of new markers was stopped after the analysis of initial data indicated that it is unlikely to result in substantial data improvement. Only in case of selection 2 of the backcross all available markers were used.

The assay position refers to the cumulative position on the genomic sequence as described in section 3.3.8. The biotinylated primer is marked in bold. The assays designed by me as a part of this PhD project are marked in green. .C – cross selection (chapter 3, section 3.4.3), PS – pre-selection of the cross material before the backcross (Chapter 4, section 4.4.1.), BC1- backcross selection 1 (Chapter 4, section 4.4.4), BC2 – backcross selection 2 (Chapter 4, section 4.4.5).

Chromo some.	Assay name	Position	Primers sequences			Sequencing	Used in			
			Forward	Reverse	C		PS	BC1	BC2	
1	pcpf06-0142	126	TCGTCTAATGCGCTTCTATCTTT	AAAAACCAAGCATGCCACAT	TCTATCTTTTCAGTGCCTTA	x	x	x	x	
1	pcpf06-0317	277	AGCAATAAAAAACGTTGCATAACAG	TTGCGGGATTTGATGTCGTATAA	ACAGAAATATATATGATGGA	x	x	x	x	
1	pcpf06-0600	510	TGAATTTAGCATAATCATAATTGG	TGCTTTACAAAGTTCAGAATGTTT	AGTTCAGAAATGTTTATCTAA	x	x	x	x	
2	pcpf01-0406	827	AGGCGGTCAGTCAAAAT	ACTAATTTCTTCTGCGGCTCCTA	ATTGGAGGGTTCCTAAAAGAG	x	x	x	x	
2	pcpf01-0323	893	ACGTCTTTTCTATGCAGACACACA	TTTCTTCCCTTTCATAAAAATCG	CTATGCAGACACACAAAGTAG	x	x	x	x	
2	pcpf01-0265	945	TAGCTGTGTGCTGCAATTCTCGAA	TTGACCATTCGCTCGTGTCTCTAAA	TCTGGTTCCTAAAAGCTAA				x	
2	pcpf01-0197	1,002	TTGAACAATTAGGAGGATCAGAA	CGAGCATTTGTATTATTGTTTCC	TGCAGGAACATGAAATG				x	
2	pcpf01-0158	1,032	ATAATAGCAATCCAAATGGACAAG	TGTTGCTCTTCAGTGATATATTCA	TTCAAAAAATTTAGAAAGCA	x	x	x	x	
2	pcpf01-0150	1,043	GAATTCGGATATTTTCGATAGTTG	TGAAAAGGGCCATAAATCATACAT	GCCATAAATCATACATCCA	x	x	x	x	
2	pcpf07-1006	1,066	TGACTCTGATAAAAGGGAAGTTT	CAAGATGTGTGTTTTGGGAGTTA	TACTTTTGAAAATGAAATTGT	x	x	x	x	
2	pcpf07-1151	1,194	CGGAATCAAAATTATAGCGAAATTGC	ATCTCCTGGAATAAAATGTCTTGC	AGTAAAAATTAGAAAAACT	x	x	x	x	
2	pcpf07-1210	1,242	TCGCTTTTCAATTTATTGTTGTCA	AAAGGCGCATATACTAGTGAATGCT	TGTTGTCATCAAGGGA	x	x	x	x	
3	pcpf02-0278	1,640	TGCCGACCCCAAAAAGTTATCCT	TTCCGCAAAAACTGTAATTCACATG	CATGGTACTAAITGTATCGC	x	x	x	x	
3	pcpf02-0452	1,777	CAGAAATGAAAAATAAAAAATGCAAT	TGTCTATAGTCTTATCATGCGATA	TGAAAAATAAAAAATGCAATA	x	x	x	x	
3	pcpf02-0675	1,949	GGAAATGTAAACAAAAGCATGTTAAT	GCTCGGAAATAATGAAAATAATTGTC	ATATGAAATAAATGTCTTTT	x	x	x	x	
4	pcpf03-0295	2,333	TCCCTTTTGCTACATCAAGATGAA	TCCATATGATTGCATACAGTAGCT	GCATACAGTAGCTCCGAC				x	
4	pcpf03-0419	2,444	CCTTTTAAAAATCGAAAAAGAAGTGA	CCGTTGCTTTATATACACAACCA	TTGCTTTATATACACAACCA	x	x	x	x	
4	pcpf03-0560	2,560	ATTTTCCTCATCAATCGAAAACAT	GGAATGTGGGGACTAGCAAAA	GTTTAGGCACTAAGACAGAC				x	
4	pcpf09-0141	2,831	AATATTTGAATCTGCACCATACCA	TTGCACAAAAAGGCTATTGTTCTC	CACCATACCATGTTTTTC	x	x	x	x	
4	pcpf09-0230	2,753	AACTGCGATGTAAGATATGTCCC	ATCAACCAATTAATAAGGTATTCOA	ATTAGAAAATAAATTATGAAA	x	x	x	x	
5	pcpf10-0697	3,099	TCCCTTCTATGTGTCTCAAAACAAA	TGGTGCTTCATCATGTTTAAATGT	ATATAATTTCAAATATGCTGG		x	x		
5	pcpf10-0868	3,247	GGTACCATATATTTCCAGATTTAA	CCCCGTTTAAATATTTACAGATAGT	TCTTTTGATTTATCAGAAAGT	x	x	x	x	

Chromo some.	Assay name	Position	Primers sequences			Sequencing	Used in			
			Forward	Reverse			C	PS	BC1	BC2
5	pcpf0-1036	3,389	TACTTCATTCGCTCCCGAAAAATA	TGTTTGATGTTGCTCTAAAGAGATG		TTGATGTTGCTCTAAAGAGAT	x	x	x	x
5	pcpf0-1203	3,532	GACGAACACAGTTTTTACAACAATG	CTTCGACCATCATGAATAGAAATGT		ATATGATTTCTGATAGTAAAT	x	x	x	
5	pcpf0-1378	3,683	ATTAGCAAAACAAGCAGGATGTAAA	CTAGTACCAAAAAATTGCCCATTTGT		TGATATGGGAAGGCAA	x	x	x	x
5	pcpf04-1072	3,864	TGGACCACTCTGGTGATAACAAG	GACCCATCACTAAACATCACTTGTA		TAAAGGAAATGGAAATTTA	x	x	x	x
6	pcpf12-0102	4,131	ATTAGCAGAAAAGGGAGAGACAA	TGGTATATCACAAAAAGGATCAAA		TGATTTTATTTCAAAGTGGG				x
6	pcpf12-0259	4,266	AGAAAAGTATGAAACAACCTTGGATT	CTTCATTGGAAACATTATTTTGAC		ACATTATTTTGACAAATTAGG	x	x	x	x
6	pcpf12-0390	4,374	CTCGTGCTATTTTTGAAAAGCAITTT	AAACATGTCGTATAACGCGAAATT		CTCAAAAAAAATCCGCT	x	x	x	x
6	pcpf14-1522	4,529	TTCATCCGGTTTTTTTATTTTCCTAT	GCAAAAAAATGTAGGGAATTTCA		AATTCGGAAAAATTGAAA		x	x	x
6	pcpf07-0962	4,928	TGAAAAAATAAGATACCCCCAAAT	TGTTTGAAAAAACTTCCAAATGAT		TGATTGATGTAATTGGGAC	x			x
7	pcpf11-1885	5,204	GGAGAAAATAGATATGAGGGGAAACA	CCTTCATAATTTTGGCATGTATAA		GTTAATATATGGCACATAGA	x			x
7	pcpf11-1467	5,570	TGTCCTAATCCACTAGCGCCTTT	GAAAGGGGAAAGAAAGCAAA		GGATATAAACATACACTGGC	x	x	x	x
7	pcpf11-1498	5,533	TTCTCCGTTTACTAGCGATGTTT	TTCCCTCTTTCCAAAATATTATCAG		AGTTAAATCCATCCATTTTC		x	x	
7	pcpf08-0431	5,812	TTGGAGGGAATATGATTATTAGTT	CATAGCGATAAAAGAAATGGAAAAA		TGCCTATATTCTGTAAATGA	x	x	x	x
7	pcpf08-0580	5,942	CCTCACAAATATCCTTTGATGACACT	TGTTTGTCCAGCCATAAAAAGC		TTGTTCAATTACTCTTTGAATT	x	x	x	x
7	pcpf04-0754	6,084	TCCCATTAGCTTCCCATATTCTT	TATGGCTTTATAAGGGGAGAAACAA		TTAGCTTCCCATATCTTAC	x	x	x	
7	pcpf04-0820	6,137	TTGGAAAATTTAAAAAGCGAGACATC	CCGCCATAAAGCTTCAATTTAATTGAT		ATTCAATAATTCAGTATACTA		x	x	x
8	pcpf07-0175	6,513	AGCATTTTCATTTTGATGTGTCITT	AACCAATTAAACCACCTATGACAAA		TTATTATCATCTTTCATCAGT	x	x	x	x
8	pcpf07-0272	6,661	CTTTTTCGGGTTTCTCAAAAATA	CAGAAACGATTATATCGAGCTTCAC		ATCAAAATCGATACCCATA	x	x	x	
8	pcpf03-0719	7,029	TTCCGCTATAAAAAACGAAGAAC	CTTTTAAATGCATTTTCACTACCT		TCACTATTTAAATCTTTTCAT				x
8	pcpf09-0600	7,282	TGTGTAGATAAAGACGAGTGGATT	CTATGTTTTTTTCCAAACCATCTTT		AACTTTATCCACAACATT	x			
8	pcpf09-0795	7,456	AACCTGGAGATGTGCGATAAAATATG	TTTCAATTCCGAGATTTCTTCATCAT		TTTGTTCAATTTTGGTGAG		x	x	
8	pcpf09-0903	7,548	ATATTGATCGCAAAATTTTATCGG	ACAAAAATAAAAACAATGGACCA		GGATTCCGAAAGTAAATAAAA	x			
8	pcpf09-0944	7,585	GAAAAACGAAACAAATCAACATAC	TCCAGTAAGACTTGTGGCTATAGG		GCTATAGGCTTATTAAATTTG				x
8	pcpf09-1203	7,810	CGAATTAATGCACGCAATAAA	TTTCACAATAGTTATTGACACACA		CACGCAATAAAATTTTACTA	x	x	x	x

Chromo some.	Assay name	Position	Primers sequences			Used in			
			Forward	Reverse	Sequencing	C	PS	BC1	BC2
9	pcpf1-0234	8,285	GGATGGATATATGGTATACCTCGA	TGATGTC AAAGGTCGATTGTT	GTATACTTCGACCATCATG				x
9	pcpf1-0377	8,408	TTGGAAGCTCATCCATAATAAAG	AGAAAGAAATGGAAAGAA GATTTC AA	GCTCATCCATAATAAAGG	x			x
9	pcpf1-0651	8,649	GAAGCAGGTTT GGAAAGCAGAAAT	GTTCACATGCATTCGGAAATAAC	CGATCATATGTTTGCC	x	x	x	x
9	pcpf1-1100	9,026	AGCAGGAATAGCTGTCGAAAATA	GGTACAA TGAGCTGCTCCAATTT	AATTATGTAACATCCGTTAG	x	x	x	x
10	pcpf04-0290	9,876	GACAAAATTGCACATTTGATGG	GCAAA CCATGCTTCA TC AAATAA	ATTGCACATTTGATGG	x	x	x	x
10	pcpf14-1029	10,423	TTTAGATCAGTAGGTTT ACCTTGG	CATGCAAAAAATGTCCAGATTC	TTACATCAACCTAGTAGTCA				x
10	pcpf14-0920	10,505	AGGAAGATCAAAATTCGCTAGTTAT	TACAT CCAAAAATAGCTTCCTCTCT	TGGAGTGATTATTTTGA				x
10	pcpf14-0562	10,807	TCGCCCTATATGACGGACTTTGA	CAGGTGATCGTTT TGTACGCAATA	TCAGTTTGTAAAGCTAAAAAT	x	x	x	x
10	pcpf14-0102	11,191	ATGGAAACCTTGCAATCATATTA	CAGGACA AAAAATAAAAAATCA TTAA	AAACGTATTTTACCTTTT				x
11	pcpf05-0177	11,583	GTGATCTTT AGCTGCTTGATCTTC	GGAGAAACGATCTTAAAGCTATG	TTTAAAGATGCTTTAAACA				x
11	pcpf05-0425	11,787	AACTATTGTCACATGGGGTTCATT	CAGATTTT TGCACCGTTTCAATAT	AAAAATGAAAAATGCAGC				x
11	pcpf05-0585	11,924	GCTGATGTTT CAATAAATGGTATG	AGAGGAGCCGATGGAATAATAATA	ATAAATACACAAATGAAGAT	x	x	x	x
11	pcpf06-0782	12,027	ACTGGTTGAAAAATTTATATGAAGA	GTTTCATA TTTTTGGAAAAATCAAT	AATATGTTTAAATGCTTGTTT				x
11	pcpf06-0902	12,564	GTATGTA CACCCATG TG GTGATTT	CGAAACGGAAACCATATATATGCC	GAACCATATATATGCCATC	x			x
11	pcpf06-1001	13,088	TGCTTAGGCATA TTTGCTTCAA	TTGAAAGGTTTGTAGCTCCTTCGA	CCATAAAATTAATTTCTGG	x	x	x	x
11	pcpf06-1044	13,122	TCGTTTACCTAACCCCATTTCA	GAAA GGAGGAAAAAATGAT CT GG	AAAAATGATTTTAAATAACCA	x	x	x	x
11	pcpf06-1102	13,291	TTGGAAGCTCATCCATAAAATAAAG	AGAAA GAATGGAAAGAA GATT TCAA	GCTCATCCATAAAATAAAGG	x	x	x	x
11	pcpf06-1154	13,333	TCTTTTTTTTCGAGGACGATA	TGGA AAAAGAGTTT TTT AAAAATATGC	AATACTAAAAATTTAATACA	x	x	x	x
11	pcpf06-1305	13,468	GTGAACAGCTTTTAAATGGCAGAA	AAAT CATCCCTATAAAATCC CACAT	CATTGTGGAGGGAA	x	x	x	x
11	pcpf06-1338	13,489	GTTGTTGGGTTGCTGAAAAATAA	TGCAAC ACAGGTA GATT ATTAGA	ATCCTATTGAAAGTAACCTA	x	x	x	x
11	pcpf13-2057	13,502	AATA AAAGAAAAAG TTT TG TG T	CATCTTTTCAGCAGAAAAATTCAT	GCAGAAAAATTCATTATTAAA				x
11	pcpf13-2207	13,637	ATATGCTCT CCAGGAGGTAT TACA	AACCCCATGCTCACAAAAGAT	TTTTTTTCAAGGGTATATAA	x	x	x	x
11	pcpf13-2508	13,909	TGTGCCATTTTAAATTTGATCTGA	GAAAA GTATACCCCAAT TGA AGCT	ATATTAGCACCGTTAAAAAT	x	x	x	x
11	pcpf13-2620	14,023	TCGGATTTT TATGTA AA GGGCTAT	CAATTCTGCAAAATTAATTAGAAACA	CTGC AAAATTAATTAGAAACA	x			x

Chromo some.	Assay name	Position	Primers sequences			Used in			
			Forward	Reverse	Sequencing	C	PS	BC1	BC2
12	pcpf10-0317	14,435	AAATTCAATAGATGCAACCAACTCA	TGGACATGATTTTTAAATCGTTCA	TTTAACTGCTTCTAAACCTA	x	x	x	x
12	pcpf10-0435	14,553	CTCATGTAAATTATAAGCTCATCA	GGCAATTTCTCTCAAAACATTCCT	ACAACAATTATGGGCC				x
12	pcpf10-0574	14,678	GGGAAGTTAAAAATTGCTGATTTTG	CATTAATAATTCGGGGGGCTCTA	ATCTTCTAAATATGCATTTG	x	x	x	x
12	pcpf03-0868	14,861	GTCGCAAAACACATCGAAATC	CATCGGCAAGCACCAATTT	TCGTGTTTTTCATGATTT	x	x	x	x
12	pcpf08-1006	15,102	CGACCAAAACAAACAGTAACAGA	TCCCCTATATGGATTAACTTGAAA	TCTACTTTGCCTGGACA	x	x	x	x
12	pcpf05-0696	15,380	CATTTCAGGACTTTCCTTCATTAC	ACATGAAAAATGCGAAATGCTTTTAA	ATGGATGAAACAAAATGATA				x
12	pcpf05-0991	15,639	ATGACCCCAATAAATTTTCAAGAA	AACAATTTTCCTTTTCGGTATCTG	CCCATAAAATTTTCAAGAAG				x
12	pcpf05-1178	15,783	GAAAAAACAAAATCATTTGGAAGAA	ACATGTTTTGGGGCTATGGTA	TTATAAAATGAAAGGCCA				x
12	pcpf05-1260	15,859	AAGAACCCGTTTTTAAAAATCGTCA	TGTTTCACTTTTCATCCTCCATAA	TCAAAAAGAAAAAAAATAATAC				x
13	pcpf14-1658	16,348	GTTTGGCTCAATTTCATCTTTAT	GGTATATCGAAAAATGTTGTAGGTA	AATATCAGCAAAACATATCAT				x
13	pcpf14-1975	16,633	AGTTGGAGACCCAAACATCATAT	TGAAATATCTGCACTCCATACACT	TATTATGATATAAACGATCC	x	x	x	x
13	pcpf14-2118	16,739	CCGATGAAAAATTACGTAAGTTTG	GGTTTGGTTGGTAAGGCTAAAT	AAAAAAAAGTTGTAGAGC				x
13	pcpf14-2445	17,047	TAACGATACCCAAACATCATCAAA	TGCACATGTTTCAAAATTTTATCTC	AATATTAITCAATTTTCACTT	x	x	x	x
13	pcpf14-2708	17,327	AGGCAACGAAGAGCTGTTAATA	TGGCATTTCGTCCACATATATCTA	TGCCACTTTTAACTGATAC				x
13	pcpf13-0906	17,575	TCAAAAAGGTTTTTGGGTTTCATACA	ACCAGCTACTTCCCAATCCCAATCCAT	CTTCTGAAGCTTTTATGAAA				x
13	pcpf13-1050	17,686	AGATTTCAGATGGAGAAAGTATCA	TTTTCTATCTTGACGTGCATAACC	GATTTTCATTAACTGGTGG	x	x	x	x
13	pcpf13-1356	17,962	TTGAAACTGCTTCAGATTGTCTT	CATTACCCAACTTAAGCCATTACT	ATCATCTTTACATATGCTTG				x
13	pcpf13-1507	18,062	ATTTGGATGATCCATTGTGTTTC	AAAAAACAGGTCAAGCAAAAAGC	TGATCCATTGTGTTTCATAA				x
13	pcpf13-1800	18,370	GACAGGGTTAATTGTATGGG	CAGGCCAATAAAAAACAGAGAACATA	CATTTTAATCCTGCAGC		x	x	x
13	pcpf13-1956	18,507	TCGCTTGTTCCTTTCCAAATCAG	TTTGAGCCAGAAAAACAGGAAGAG	TCCAATCAGATTGAAAAT	x			x
14	pcpf13-0185	18,900	CACAAATTAATGGTGGCAGAGAT	TTGGAGGTAATGTTTTATAAGGG	ATAAGGGAATATTTATCTAC				x
14	pcpf13-0448	19,144	TGATATTGGTGGATGTAAAGAACAA	CGAGCTGTAAATGTTTTTCCCTGTA	CCCAAGGGTGTTTTACT	x	x	x	x
14	pcpf13-0737	19,373	ACCCCGAATATCTTCTCTTATTGA	AATCGAATATGCGCAAGGCTTA	TGTTGTTTCATGAGAAAAA				x
14	pcpf07-0524	19,515	TGACGGGAAAAAAAATAATGTCCTG	GGCCAATGGCTAGTGTCAATA	CAATGGCTAGTGCATAAAT	x			x

Chromo some.	Assay name	Position	Primers sequences			Sequencing	Used in			
			Forward	Reverse			C	PS	BC1	BC2
14	pcpf08-0650	19,605	CAAGTCTCAAAGCGCACAAAT	TGGTTTTTTTGATCACCCCTTTCAAT		CTATCGAGGCAGTAGTAAA	x			x
14	pcpf08-0819	19,758	GAAAGCTATCCGCAACCTATTAGA	TTCGACTGATGATTTTACAAAGTCT		TTATGTCTAATTGATTTTGG				x
14	pcpf12-0627	19,886	GGTGGACCATCTTCTTCTAAACTT	TGCTTTATATAATCAAGGCCCTCA		CCATACCTTTTCAAAACAGT	x	x	x	x
14	pcpf12-0934	20,146	ATGCTGCTGCTATGAGATATACA	AACATCAACAAGGTTAIGACTTGG		GAAAAAGAACCTAAAGTG	x	x	x	x
14	pcpf12-1538	20,679	TTTTTCCTGGCTTAATTTATCAA	TATGCCCCAAGCCAATATTGTT		AATTTATCAATATTATGAGG	x	x	x	x
14	pcpf12-1682	20,798	TAA CATCATTTGCGATTAAATGAAA	GGCAGATGAATAATACTTGTGCGAT		TTCTGTTAATACTGGTTTIG				x
14	pcpf12-1830	20,874	ATCTGCACTTTTCCCTGACTACTAGC	AAGCAGAAGCCTTTTATACACTGGT		TATGACAGTGGAAATCCT				x
14	pcpf12-2125	21,127	GAAATGAAGGTGCTAATTTTGAAA	TCGTTGCA TTCACCTGATTT		ATTAGAAAAAGTTTGTGC				x

Appendix D – The alignment of the three conserved genes connected to CQ/ART resistance.

All alignments according to Clustal W2 software. The symbols indicate:

*- a identical position

: - a conserved position

. - a semi-conserved position

Aat1

P.falciparum_PFF1430c	-----MN-KKYGTSSNNHNN--KKDKKNN	21
P.vivax_PVX_114575	-----MNENGVGENKNNFKKKEKRSKKG	24
P.knowlesi_PKH_111940	MLILTAQHAPPPTPCWNPLOFSYFYKMNENGI IENKNNSKKKEKRSKKG	50
P.chabaudi_PCHAS_112780	-----MVNNIDTLDYCHSANDVVSNIVN	23
P.berghei_PBANKA_112830	-----MANNIDILDYCHSTNDVVQNIVN	23
P.yoelii_PY04168	-----MVNNIDILDYCHSTNDVVPNIIVN	23
	: . . :	*
P.falciparum_PFF1430c	ADKNKNKKNTTTTEENKDSNKSLVNNDSSKNDSSK--NKYNIVKANIKNI	69
P.vivax_PVX_114575	PKKAGSKDDHAKS DPAGAPSGEVNSGNAREGDEAKENSAHPPSKNNMNSNL	74
P.knowlesi_PKH_111940	PKKVDSKEDHTNSDPAGTPSGNVNNGNTPQVEEKKENNTASNKNVSNTL	100
P.chabaudi_PCHAS_112780	YKK--NNDNYNYRGKYDDGNTGISKNMG-----ANSNYNA	58
P.berghei_PBANKA_112830	YKK--NNDNYNYLEKHDDENAGISKNNIG-----VNSNYNRV	58
P.yoelii_PY04168	YKK--NNDSYNYLEKHDDVNAGISKNNMG-----ANSNYNGV	58
	. * . . . : . . . :	.
P.falciparum_PFF1430c	FASDKKNEKSDK-----NEKN-ESSKSSKNTETTYTNVNDKKSNNLITKG	112
P.vivax_PVX_114575	FSSNKKKKNQKKKYESNMGENNADMNGNKLKTS PYVNLGGDDDMNEG-AEG	123
P.knowlesi_PKH_111940	FSWKKKKKNQKK-CENKMEENNDSIHGNSLKTS PYVNLGGDDDINEVGEEG	149
P.chabaudi_PCHAS_112780	FMFNKKIGNEAKQIDTNMENITPYINIDSLENTTKINLIGSDKDNMRD	108
P.berghei_PBANKA_112830	FVFNNKRGDETQKIDTNMESITPYININSLESTKKINLVDESSKNYMMGN	108
P.yoelii_PY04168	FVFNNKIGDETQKIDTNMENTTPYININSLENTSKIDLIDESSKNYMMGN	108
	* . :* .. * . . : . . . : *: * . : :	.
P.falciparum_PFF1430c	SN---DKKKKKKDSKK-----NSSNNNNNNNTIVDI	139
P.vivax_PVX_114575	DGYYGEDNNNGAKDVQKNCQEGR-----ENENCDCDNQNSTIVDV	163
P.knowlesi_PKH_111940	DGYYYQDNNNDVNDTQENYQESN-----ATENYDENAHENAIADV	149
P.chabaudi_PCHAS_112780	HWFHNKIKNEAKNNKNNSKNKN-----ISNQYVSLAHENVNDQEI	189
P.berghei_PBANKA_112830	NWYHQNKYNNKMNSKNSKN-----SHNKHISIAHIKYDDQEI	146
P.yoelii_PY04168	NWYHQNKYNNKMNSKNSKNSKNSKSSKNSKSHNKHISIDHIKDDQEI	158
	. : : : : . : : : :	:
P.falciparum_PFF1430c	SDGD-YTNDEEGTNPKRNWKGRF FSRFTPGGVRSSTVLFLCTAIGVGFL	188
P.vivax_PVX_114575	GENEKNPEDEEAEEKS-KKNWQRFR FSRFTPGGVRSSTVLFLCTAIGVGFL	212
P.knowlesi_PKH_111940	GGNEKNAQDKKEEKRRKNWKRFR FSRFTPGGVRSSTVLFLCTAIGVGFL	239
P.chabaudi_PCHAS_112780	NEGNKKLCEENETNTKKTWQKRAF GPFTPGGVRSSTVLFLCTAIGVGLL	199
P.berghei_PBANKA_112830	NEKGKNKLYEENQTNTKKTWKRRA FSPFTPGGVRSSTVLFLCTAIGVGLL	196
P.yoelii_PY04168	NEKEKNKIYEENQTNTKKTWKRRA FSPFTPGGVRSSTVLFLCTAIGVGLL	208
	. : * .. *: *: *: *. *****.*:*****:*	:
P.falciparum_PFF1430c	SIPYVFSKLGIIILSIILIILNAFESYVTTNILCTSSEHNTFVYGNLLKK	238
P.vivax_PVX_114575	SIPYVFSKLGIIILSLLLILLNAVESYVTTNILCLSSLEHNTFVYGNLLKK	262
P.knowlesi_PKH_111940	SIPYVFLKGIILSLLLILLNGMESYVTTNILCMTSLEHNTFVYGNLLKK	289
P.chabaudi_PCHAS_112780	SIPYVFSQLGIIILSIILILLNALESYITTNILCMSALEHNIFVYGNLLEK	249
P.berghei_PBANKA_112830	SIPYVFSELGIIILSIILILLNSLESYITTNILCMSSEHNTFVYGNLLEK	246
P.yoelii_PY04168	SIPYVFSSELGIIILSIILILLNSLESYITTNILCMSSEHNTFVYGNLLEK	258
	*****.*:*****.*:*****.*:*****.*:*****.*:*****.*:	:

P.falciparum_PFF1430c	IGNKYYKTIIDFGLSFGFVSSYILILILISNFLTSTIFYVFNFPFTLTNNV	288
P.vivax_PVX_114575	IGHKYHKTIIDIGLTFGFLSSYILILILISNFLTSSILYVDFPFSFLYNHT	312
P.knowlesi_PKH_111940	IGHKYHKTIIDIGLTFGFLSSYILVLILISNFLTSSILYVFNFPFSFLYNHN	339
P.chabaudi_PCHAS_112780	IGNKYYKTLIDLGLTFSSFLSGYVLILILMNDYLSNVLHTFNFPFSFICNRV	299
P.berghei_PBANKA_112830	IGNKYYKTLIDFGLTFSSFLSGYVLVLILVNDYLSNILYTFNFPFSFISNRI	296
P.yoelii_PY04168	IGNKYYKTLIDFGLTFSSFLSGYVLVLILVNDYLSNILYAFNFPFSFISNRI	308
	**:*	
P.falciparum_PFF1430c	FLVLICLLILPITFRNKVGSLSNHLFISLSLSITVLTIGLQTKSYNNL	338
P.vivax_PVX_114575	FLIILVCVLIVPITFRDQVGSLSNHLFISLSLSITVLTIGWQARDYYTI	362
P.knowlesi_PKH_111940	VLIIIVCSLIVPITLRDQVGSLSNHLFISLSLSITVLTIGWQARGYYTV	389
P.chabaudi_PCHAS_112780	FITIVICLLVPLTFRHIGSINYFLIFSLSITLTVLTGVYQSKYYMSL	349
P.berghei_PBANKA_112830	FITIVICLLVPLTFRHIGSINCFVFSLSITLTVLAVGYQSKYYMSL	346
P.yoelii_PY04168	FITIVICLLVPLTFRHIGSINCFVFSLSITLTVLTGVYQSKYYMSL	358
	.: *:*:* *:	
P.falciparum_PFF1430c	LINKEVNLFKMDKHFFKCFNILLFSFSQQPNACFITGQFNQPTHRLNKS	388
P.vivax_PVX_114575	LINKEIVLFNMDKHFFKCFNILLFSFSQQPNACFIAGQFNQPTHRLTKL	412
P.knowlesi_PKH_111940	LINKDIVLFNMDRHFFKCFNILLFSFSQQPNACFITGQFNQPTHRLTKS	439
P.chabaudi_PCHAS_112780	LPEKNISLKFINKHFFKCFNILLFSFSQQSNACFITGQFNQPTQRRITKS	399
P.berghei_PBANKA_112830	LPEKNISLKFINKHFFKCFNILLFSFSQQSNACFITGQFNQPTQRRVTKS	396
P.yoelii_PY04168	LPEKNISLKFINKHFFKCFNILLFSFSQQSNACFITGQFNQPTQRRITKS	408
	* :*:* *:	
P.falciparum_PFF1430c	TFRSVILQVIFYTYLFGILGYFSFLNTAKDNIVLNYENSNVSILLCKFLLS	438
P.vivax_PVX_114575	AYRSVLLQIIFYTYLFGLLGYLSFLTTKDNIILNYEDTNVSILLCKFLLS	462
P.knowlesi_PKH_111940	AYRSVVLQIIFYTYLFGLLGYLSFLTTKDNIILNYEDTNISILLCKFLLS	489
P.chabaudi_PCHAS_112780	ESRSILIQIIFYTYLFGLLGYLSFLNTTKDNIILNYEETNMSILLCKFFLC	449
P.berghei_PBANKA_112830	ESRSILIQIIFYTYLFGLLGYLSFLNTTKDNIILNYEETNMSILLCKFFLC	446
P.yoelii_PY04168	ESRSILIQIIFYTYLFGLLGYLSFLNTTKDNIILNYEETNMSILLCKFFLC	458
	**:	
P.falciparum_PFF1430c	LITFFFSVPLNFMGYSQSMALGITTDRDALYKLYTYIFRRTGYSANLSLL	488
P.vivax_PVX_114575	ITFFFSVPLNFMGYSYSSLLSLYLSARNMFLKFYVVFVTRNRNRYMANLSTFL	512
P.knowlesi_PKH_111940	ITFFFSVPLNFMGYSYSSLLSLYLSARNMFLKFYVFSRRNRNRYMANLSTFL	539
P.chabaudi_PCHAS_112780	ISFFFSIPLNFIATYPSMMSLYTTIHNNIQKLYSALFTRNEYLPSPSLNIL	499
P.berghei_PBANKA_112830	ISFFFSIPLNFIATYPSMISLYTSIRNKIQKLYSVLFTRNEYLPSPSNIL	496
P.yoelii_PY04168	ISFFFSIPLNFIATYPSMISLYTSIRNNIQKLYSVVFTRNEYLPSPSNIL	508
	:*:	
P.falciparum_PFF1430c	SEYTND-----PYQETHADNITEHSSSVSESQTDQNRQMWISVIVTIFCA	533
P.vivax_PVX_114575	REDEANPFEEAPPFEENALDNLTESSSTTESQAEDKTQRKLISISVTILCA	562
P.knowlesi_PKH_111940	REDEDN-----IFEDNALDNLTESSSTTESQAEDRKQRRLLISICVTILCA	584
P.chabaudi_PCHAS_112780	RYDTEN-----PFGEYAQDGNTENSRAVESQGDDMFQRKFVAIFVTCLCA	544
P.berghei_PBANKA_112830	RYDTEN-----PFDEYTLDENTENSSTSESQGDDMFQRKCAAFVTCLCA	541
P.yoelii_PY04168	RYDTEN-----PFGEYIQDENTENSSESQGDDMFQRKCAAFVTCLCA	553
	: : : * *:*:* :*:*:* :* * :* * :* * :*	
P.falciparum_PFF1430c	LIACKVKKLSNVIGIGGGITSTLISCLLPNLIYYKNRHNVSNKLRYSTL	583
P.vivax_PVX_114575	LIAFNVKKLSNVIGIGGGITSTLISW-----	588
P.knowlesi_PKH_111940	FIAFNVKKLSNVIGIGGGITSTLISCLLPNLIYFKNRHNVNKNKLERYSTL	634
P.chabaudi_PCHAS_112780	FVEFNVSKLSNFIGIFGGFTSSIIISCLLPNLIYYKNMHTFKNKTERYATL	594
P.berghei_PBANKA_112830	FVEFNVSKLSNFIGIFGGFTSSIIISCLLPNLIYYKNMHTFKNKIERYATL	591
P.yoelii_PY04168	FVEFNVKRLSNFIGIFGGFTSSIIISCLLPNLIYYKNMHTFKNKTERYATL	603
	: : : * *:*:*:* *:*:*:*:	
P.falciparum_PFF1430c	FMLCFFSFMGFLSVVVTTLNLIL-----	606
P.vivax_PVX_114575	-----	
P.knowlesi_PKH_111940	CMLCFFSFMGFFSVITTALILIF-----	657
P.chabaudi_PCHAS_112780	ALLCFFSAMGLISSIVTAFIIIR-----	617
P.berghei_PBANKA_112830	ALLCFFSVIGLISSIVTAFIIY-----	614
P.yoelii_PY04168	ALLCFFSVMGLISSIVTAFIILCIYMHSLDFNVIILNDLHI	645

P.falci	parum_PFB0675w	MLELLLRNAYKKEYEYILGLSSKSQIYIKINILNYCLLFSFFLFYASYFFFFI	50
P.vivax_PVX_002795		MMPFLSNTYSKCREQIARKLTISSKRQIYIINNLSLFSFILFYISYFFFFI	50
P.knowlesi_PKH_040510		-MASLNNIYSKCREHIARFTTISNRKKIYIINNLSLFSFILFYISYFFFFI	49
P.chabaudi_PCHAS_031370		-MVFFYSIYSALMKYMTNKLSYSRKQKIFVINNCFFSFILFYISYFFFFI	49
P.berghei_PBANKA_031160		-MVFFYSIYSALMKYITNKLSYSRKQKIFVINNCFFSFILFYISYFFFFI	49
P.yoelii_PY05194		-----	
P.falci	parum_PFB0675w	NGIINPAVKIVFHGSIIE D KTKVYENSIIESIWLFLTGSYFNMAMLLFF	100
P.vivax_PVX_002795		NGIINPAVKIVFHGNIADNTQVYENSILESIWLFLTGSYFNMAMLLFF	100
P.knowlesi_PKH_040510		NGIINPAVKIVFHGNIADTTQVYENSILESIWLFLTGSYFNMGMMLFF	99
P.chabaudi_PCHAS_031370		NGLIYPVKIVFHGSVIANQTIVYKNSILESIWLLYTSGSYFNMGVLLIF	99
P.berghei_PBANKA_031160		NGLIYPVKIVFHGSVIANQTIVYKNSILESIWLLYASGSYFNMGVLLIF	99
P.yoelii_PY05194		-----MGVLLIF	7
		*.:**.*	
P.falci	parum_PFB0675w	SIIPFLKLIMVSDNFYSFIVLYKMNNKKREEEKRRRRIRAREYNNAKY	150
P.vivax_PVX_002795		SIIPILKLIMVSDNFYCFKLYNLSLQHEEEEKEEE-----	137
P.knowlesi_PKH_040510		SIIPILKFIMVSDNFYCFKLYNLSLQHEEEEKEEE-----	136
P.chabaudi_PCHAS_031370		SVIPIILKFVMVTDNFYSFFKLYNLGLKFENENDKDD-----	135
P.berghei_PBANKA_031160		SVIIPVLKFVMVSDNFYSFFKLYNLSLNFNENDKDD-----	135
P.yoelii_PY05194		SVIIPVLKFIMVSDNFYSFFKLYNLSLNINENDKDD-----	43
		*.:***.**:**.***:*. **:. . :*:..	
P.falci	parum_PFB0675w	NKNRYKMNKYNRNKNINDIYKDDIYYSENIFKNDEINYINTINENEFEFIL	200
P.vivax_PVX_002795		-----EDDIFFINTVNENEQFIL	155
P.knowlesi_PKH_040510		-----EDDIFFINTVNENEQFIL	154
P.chabaudi_PCHAS_031370		-----AKLIYYINIINVNEEFAL	153
P.berghei_PBANKA_031160		-----AKLIYYINIINVNVGEFAL	153
P.yoelii_PY05194		-----AKLIYYINIINENEFEFAL	61
		. * : * * : * * : * *	
P.falci	parum_PFB0675w	KKFKILNFIISRQFQVDVFISLFI VSSLNLYLLEARMLNGAYYFLNYCMLS	250
P.vivax_PVX_002795		KKFSILSSISRQFQVDVLISLFI VSSLNLYLLEAKILSGAYHFLNYCLLS	205
P.knowlesi_PKH_040510		KKFSILSSISRQFQVDVLISLFI VSSLNLYLLEAKILSGAYHFLNYCLLS	204
P.chabaudi_PCHAS_031370		KKFSILSSISRQFQVDVFISLFI VSSLNLYLLEAKILIGAYHFLNYCVLS	203
P.berghei_PBANKA_031160		KKFSILSSISRQFQVDVFISLFI VSSLNLYLLEAKILIGAYHFLNYCVLS	203
P.yoelii_PY05194		KKFSILSSISRQFQVDVFISLFI VSSLNLYLLEAKILIGAYHFLNYCILS	111
		.*. **:*:*****:**	
P.falci	parum_PFB0675w	TISSFFLLFSFTSLKIHFKNGNIKISACLNESNLEVTTSGPLSTKDVLVEE	300
P.vivax_PVX_002795		TISSFFLLTLTSLKIRIFKRGNIKIYSLPAERRLSVRTSLQQENTRESR	254
P.knowlesi_PKH_040510		TISSFFLLTLTSLKIRIFKRSNIKIALPTECLSVRNSLQ-EDARDNR-	252
P.chabaudi_PCHAS_031370		TISSYLLLTFMKIKINIFKEGNIKIFSSRNENSVAGKNSDSMLCNGKRD-	252
P.berghei_PBANKA_031160		TISSYLLLTFMKIKINIFKDGNI RIFSSRGENILTGNKSGTIIYENENK--	251
P.yoelii_PY05194		TISSYLLLTFIKIKINIFKDGNI RIFSSRDENIANGKNTDXIIYDNK--	159
		****:~::~. ~:~.*** ~:~.* : ~: ~:	
P.falci	parum_PFB0675w	EGHAQINNLIINDKNMTSGVVNDMSGNGNVELTDDLKSGEPQNRRDIQT	350
P.vivax_PVX_002795		-----DLDCNTNGEGVALPGQLRNGEENDTLTGEPANKMSVTRG	292
P.knowlesi_PKH_040510		-----MLDCVGEGEVFGHGKLGKNGK--NDSPVVEV NKIDLKRG	288
P.chabaudi_PCHAS_031370		-----ILKKESSK-----IEFNKIY--DSNNTKESHQWNKKN-	282
P.berghei_PBANKA_031160		-----EFNNIYDNNNNNTKDNHKWNKKN-	274
P.yoelii_PY05194		-----ELNNIX--NNNNTKENHKWNKKN	181
		: . : . : . : . :	
P.falci	parum_PFB0675w	EETKKEKMNTR-----THNDEDNTKKKNIKDKKKKANGDDKV	386
P.vivax_PVX_002795		GGDDPEEGEE-----HNRRGANHSEHLNLSARKGAEPLGNA	328
P.knowlesi_PKH_040510		GKNNLDEGEDDDDDNNNNNNNNNNNNNNNNNNNNNQLD SILRKGVEWMGLA	318
P.chabaudi_PCHAS_031370		--KKKS NKE-----DLNSDNMNVTHKKKMTSEKDD	310
P.berghei_PBANKA_031160		IKTKNSKENL-----GGLNSDNTNVAHKINNESENE	306
P.yoelii_PY05194		KKKKNSKENF-----DGLNSEN-----INNESENE	207
		. : . : . : . :	

P.falciparum_PFB0675w	IQKCIDNERKKKQNGMIQSVNDGDKNSFNNNNNNNINGDSNNNNNINGDS	436
P.vivax_PVX_002795	KNIRRGKKRVSHGRRPPREDPHQAGVTGMIDNRGVYHRDFPRASMHMMH	378
P.knowlesi_PKH_040510	KKNRKGGKKKVFHEEGSMQGD-YTSGTTGIIITHRGVYHLDPISIPRVNQRG	387
P.chabaudi_PCHAS_031370	KIGHINSKYKMDNEIKSLEEDTSAGKT-----	338
P.berghei_PBANKA_031160	KIGHMNSKYKIDSNGNKSLEEDTSAGKT-----	334
P.yoelii_PY05194	KIGNINNKYKIDGNENKSLEEDTSAGKT-----	235
	..: : : : . .	
P.falciparum_PFB0675w	NNNNINGDSNNNNINGDSNNNNINGDSNNN-----NINGDSNNNNYH	478
P.vivax_PVX_002795	HMHHMHHRGGALPRGSGSWKGEAEMPHDKMPHLREELNDNCPNERGALG	428
P.knowlesi_PKH_040510	IHSREIHFSGAMQRGSGSWKGETELITYD-----HEDLNNS-VNERGRHS	431
P.chabaudi_PCHAS_031370	-----TGDENMSPNKQ-----NDC-----	352
P.berghei_PBANKA_031160	-----TGDEHMTSKE-----NGR-----	348
P.yoelii_PY05194	-----TGDEHMTSKE-----NEC-----	249
	..: . . . *	
P.falciparum_PFB0675w	NNYHNNYNNYHNNYRNNNCRNNILEQNKCDKNVLCYNNIYNTMKDNDTY	528
P.vivax_PVX_002795	QKSGNTEPSQFPHPHERYERYEKHLLCYNKLQT-LSNINYLYKVMKNNDAY	477
P.knowlesi_PKH_040510	KKKVNLESTQLQPHER---YEKQLLCYNKLQT-LNNINYQYKVMKNNDAY	477
P.chabaudi_PCHAS_031370	EDIDSIDGENVIEHNS---YENNILCYNKLQS-LNSVSYLYKVMKNNDAY	398
P.berghei_PBANKA_031160	EEIDSIDGENVAENNS---YENNIMCYNKLQS-LNSVSYLYKVMKNNDAY	394
P.yoelii_PY05194	EEIDSIDCENIAEKNS---YENNILCYNKLQS-LNSVSYLYKIMKNNDAY	295
	:. . : : .: : ** : . : * : ** : ** :	
P.falciparum_PFB0675w	IYLKKNKFNSLLKSNCIKTNFNMIKIGYVIFLVLCLCIYELITGVECSL	578
P.vivax_PVX_002795	IYLKKKKFNSLLKTNFHHCRNLNKIKIGYACFLFVFLCLCVYLITAVECSM	527
P.knowlesi_PKH_040510	IYLKKKKFNSLLKMFHQCRLNKIKIWIYACFLFLCLCVYLITAVECSM	527
P.chabaudi_PCHAS_031370	IYLKKKKFNLLLRFSQKTRFNTIKVAHTFILLFLLCISYIFITLVETDM	448
P.berghei_PBANKA_031160	IYLKKKKFNLLLRNFQKTRFNTIKLAHAFILLFLCLCSIYFITLVETDM	444
P.yoelii_PY05194	IYLKKKKFNLLLRFSQKTRFNTIKIAHAFVLLFLCLCSIYFITIVETDM	345
	****: ** : . : * ** : . . : ** : ** : ** : ** : ** :	
P.falciparum_PFB0675w	FGIYIYLSYFNFNIEGILIDYMDMLNILKLKIKKGYIYPFFVMLPFIFPV	628
P.vivax_PVX_002795	VGIYIYLSYFNFNVEGILIDFIDMLNILKLKINRSYIYPFFVMLPFIFPV	577
P.knowlesi_PKH_040510	VGIYIYLSYFNFNVEGILIDFIDMLNILKLKINRSYIYPFFVMLPFIFPV	577
P.chabaudi_PCHAS_031370	LGIYIYLNFFNFNIESVMIDYIDMLNILKLKVNNQYVFPFFFMFPFIFPL	498
P.berghei_PBANKA_031160	LGIYIYLNFFNFNIESVMIDYIDMLNILKLKVNNYVFPFFFMFPFIFPL	494
P.yoelii_PY05194	LGIYIYLNFFNFNIESVMIDYIDMLNILELKVNNYVFPFFFMFPFIFPL	395
	.*****.:****:*. :	
P.falciparum_PFB0675w	IISMCFFLSVFFFLNMYYESFSKLYKKISELKNEFINSSSENDNVNERILVS	678
P.vivax_PVX_002795	IIGLSFFMSVLFMNLYYVNFVRGYKKMCALNDELVSPEVED-NQKMMVS	626
P.knowlesi_PKH_040510	IIGISFFLSVLFMNLYYVNFVRGYKKMCAMNDELVSPEVEE-NQKMMVS	626
P.chabaudi_PCHAS_031370	IIAVSFMLS VFFMNLYYVNFVSVWYKKITAFSDELVDPEVQN-NQKITIS	547
P.berghei_PBANKA_031160	IIAISFFLSIFFMNLYYVNFVSIWYKKITAFSDELVDPEVPN-NQKITIS	543
P.yoelii_PY05194	IIAISFILSVFFMNLYYVNFVSIWYKKITELNDELVDSEVPN-NQKITIS	444
	**.: : : : : : : : ** ** : : : : : . * : : : : : :	
P.falciparum_PFB0675w	ETSNHLCLN---ESNDKVSNTSDDFLSRNNSNISSSKSEMINSNFVNKL	725
P.vivax_PVX_002795	ERYNYLYIKKEIEKNDRLSNTSDDLTSPNNSTISSKSEDINTSFISNRM	676
P.knowlesi_PKH_040510	ERYNYLYIKKEIEKNDRLSNTSDDLTSPNNSTVSSKSEDINTSFIFSRM	676
P.chabaudi_PCHAS_031370	ERHNYLCIKKEIEKNDKLSNTSDELSSHNNSTITSSKSEDINTEFVYSGL	597
P.berghei_PBANKA_031160	ERHNYLCIKKEIEKNDKLSNTSDELSSHNNSTITSSKSEDINTEFVYNGL	593
P.yoelii_PY05194	ERHNYLCIKKEIEKNDKLSNTSDELSSHNNSTITSSKSEDINTEFVYNGL	494
	* * : * : : * : : : : : : : * * : : : : : * : : : : : :	
P.falciparum_PFB0675w	LNIFYFSFAVFFSYLGS AFLHISLGEIICIALLTIFYQIVKHTNNLNITILL	775
P.vivax_PVX_002795	LNIFYFSLSVFFTSICSF CMHISLGEIISISLLTFYQIVKHTNLLNVLVLY	726
P.knowlesi_PKH_040510	LNIFYFSLSVFFTSICSF CMHISLGEIISISLLTFYQIVKHTNLLNVLVLY	726
P.chabaudi_PCHAS_031370	LNIFYFLMSSFFASICSILLHMSLGEIIAIAIIIFYQIVKHTYNLLNVLVLY	647
P.berghei_PBANKA_031160	LNIFYFLMSSFFASLCSILLHISLGEIISIAILIFYQIVKHTYNLLNVLVLY	643
P.yoelii_PY05194	LNIFYFLMSSFFASLCSILLHISLGEIISIAILIFYQIVKHTYNLLNVLVLY	544
	***** : : ** : : * : : : : : : : : : : : : : : : : : :	
P.falciparum_PFB0675w	KSEKIKFCKFLLFILYGLLCFSINLYVNQWEEYITKLKRLKRRILLFEKN	825
P.vivax_PVX_002795	KSNTVKFCKFLLFVYVGLLCL SINLYVSQWELYVKKLKKTKKRIALFQNN	776
P.knowlesi_PKH_040510	KSSTVKFCKFLLFVYVGLLCL SINLYVSQWELYVKKLKKTKKRIALFQNN	776
P.chabaudi_PCHAS_031370	KTDKIHF SKFILFIAGLICFSINMYVAQWEKLVKKLQKIRKQIYLFESN	697
P.berghei_PBANKA_031160	KTDKIHF SKFILFISFGLICFSINMYVTQWEKIIKKLQKIRKQIFLFESN	693
P.yoelii_PY05194	KTDKIHF SKFILFISFGLICFSINMYVAQWEKLIKLLQKIRKQIFLFESN	594
	*.: :	

P.falciparum_PFB0675w	FNVGHIYHEEKTIFYHATLLFLQGLRSVKIMNMNFYEEKGIYYLSLDGYFK	1239
P.vivax_PVX_002795	YNVGHYIYHEEKTIFYHATLLFLQGISSIKIANVHFYDYGKYNLILDGYFK	1221
P.knowlesi_PKH_040510	YNIGHYIYHEEKTIFYHATLLFLQGISSIKIANIHFYEEKGKYNLILDGYFK	1192
P.chabaudi_PCHAS_031370	YEVGRIYHEEKTIFYHATLLFLQGISSINIDNIYFYDQGGKYLTFTNGYFK	1086
P.berghei_PBANKA_031160	YEIGRIYHEEKTIFYRATLLFLQGISSINIDNIYFYDQGGNYLTFTNGYFK	1073
P.yoelii_PY05194	YEVGRIYHEEKTIFYHATLLFLQGISSKKLIQ-----	1000
	::*:*****:*****: * :: :	
P.falciparum_PFB0675w	HIIGPLFLKLCIGTNFCPISTYAFLVGSKPTFSVNVAVQCNNKKPPYYMT	1289
P.vivax_PVX_002795	HIIGPLFLKLCIGTNFCPISTYAFLIGSKPTFSITMSAHCNDRKPPNYIT	1271
P.knowlesi_PKH_040510	HIIGPLFLKLCIGTNFCPISTYAFLIGSKPTFSITMSAHCNDKTPPNYIS	1242
P.chabaudi_PCHAS_031370	HIIGPLFLKLCIGSNFCPISTYAYLIGNKPTFSITIEAQCDYAAPNYIS	1136
P.berghei_PBANKA_031160	HIIGPLFLKLCIGSNFCPISTYAYLVGNKPTFSITIEAQCDYASPDYIS	1123
P.yoelii_PY05194	-----	
P.falciparum_PFB0675w	DIIIVKDLKITKIEIVKHSVDVIDNVDIKLDDVQDRVQEKVNAMLEAKKKII	1339
P.vivax_PVX_002795	DIVVKDLKITKIEIVIKHNDIENVDIQLDDVQERVQEKVNSMLAERKRFI	1321
P.knowlesi_PKH_040510	DIVVKDLKITKIEIVIKHNDIENVDIQLDDIQERVQEKVNSMLAERKRFI	1292
P.chabaudi_PCHAS_031370	NLVVSDLQITKIEIVIKHSDIENVDIKLDDIQERVQEKVNAMLADHLQII	1186
P.berghei_PBANKA_031160	NLVVSDLKITKIEIVIKHSDIENVDIQLDDIQERVQEKVNAILADHLEII	1173
P.yoelii_PY05194	-----	
P.falciparum_PFB0675w	VWKNQKYHLEGFLNYLISKNALSGFSCEPINY---	1371
P.vivax_PVX_002795	LWKSHKYNLEGFINYLISKNALAGFSCRPMNYHDV	1356
P.knowlesi_PKH_040510	LWKRHKYNLEGFINYLISKNALAGFSCKPMNYQDL	1327
P.chabaudi_PCHAS_031370	IWKNQKYDLESFVNYLISRNSLYGFSCEPYNY---	1218
P.berghei_PBANKA_031160	IWKNKKYNLESFINYLISRNSLFGFSCEPYNY---	1205
P.yoelii_PY05194	-----	

Ubp1

Note: In P.vivax two neighbouring proteins (PVX_081445 and PVX_081440) are identified as ubp1 homologues (probably as an effect of annotation error). They are most likely two fragments of the same protein. Both are shown on the alignment.

P.falciparum_PFA0220w	-----MSHINYNVEKRKS	13
P.vivax_PVX_081545	MRERIRVEADKPLPKDESDN---LFNPWLSHFEKEKKAKLSSTCSTLQS	47
P.knowlesi_PKH_020900	MRERILTRTDKPLPHDDADDGNKLCNLWLSHIENEKKKDLGITCSRQKC	50
P.chabaudi_PCHAS_020720	-----MDKNKKKN	8
P.berghei_PBANKA_020880	-----MEKKKKKH	8
P.yoelii_PY01242	-----	
P.falciparum_PFA0220w	LKKHNNNNNNNNIYNNKIDTPNIKNYDDSSKHINTNPQVLDILLSNME	63
P.vivax_PVX_081545	KRSGNHHEYSSSYDRGDSHRCNRGDGLGHWRDDANECDDATFFILHNLKE	97
P.knowlesi_PKH_020900	KKSGKHHDYSSRYDSRNEHRCYKGDPMShRGGDAECDEAASFMLHLKLE	100
P.chabaudi_PCHAS_020720	KKLEIRDPSDEINNCEGEKGGTQFDESFPFD-----NPEPNHINENHNVN	53
P.berghei_PBANKA_020880	KKREISESDNNMNCGKEKGSIQHDELYNIEQYLSNLSNNIENQNID	58
P.yoelii_PY01242	-----	
P.falciparum_PFA0220w	KDKKLRLNNYINMFDKNKNDKVTTNHPSHNIYNRKNNDTYDDQDKDEQY	113
P.vivax_PVX_081545	NTRKGCTLAGSEKGRISREEDTLFPAQFGSTDDENFLFGKQVRGDPQ	147
P.knowlesi_PKH_020900	NIRKGCSLTHGSENGKRSSRE-DALFPShLGVADDVNLLLRKKVKGDH	149
P.chabaudi_PCHAS_020720	NDDQNHVNIKKKN-----TSNSFENDILEFLPPNVDDNSDIFHANEV	95
P.berghei_PBANKA_020880	NDYQKHITIKKKNCNEKRGKNLSNSFEDNTLEFLPQNVDNSNNFQKTEI	108
P.yoelii_PY01242	-----	
P.falciparum_PFA0220w	VDTDDSFSLNNTKKKINKRDIISYDNYIFEDEDKVSSKYLEYKNDSTSHM	163
P.vivax_PVX_081545	KVVHTLRGPSQRQERTRTKAGTTKENDTSWGEGACLLSGKRRKDRQEDG	197
P.knowlesi_PKH_020900	KFIRTVRDPsLRKGSTRIKVNEATKRNDISGGKGACLLIGKSRKDRHGDG	199
P.chabaudi_PCHAS_020720	NKIEETNLNDNEQNTMNIKNKDPNNRLDPKGND-----DNIVKNVESD	138
P.berghei_PBANKA_020880	NKIEENNLNSTEAIKMNNKKKDINNRSDDKNDCKCYDNDNADVKTEND	158
P.yoelii_PY01242	-----	

P.falciparum_PFA0220w	KKKKDEGSNRKGN-----NMSNNNDNN	188
P.vivax_PVX_081545	ALRGDRHRRRHQTNGHHEELYLSDVYHSDMHVSDMYSGDGDFLRADGES	247
P.knowlesi_PKH_020900	AVRRDRHGCGHHQK-----TNEYDEELYMNYGDSCDGEILEVDGVN	240
P.chabaudi_PCHAS_020720	FEKNN-----LLVTDENV	152
P.berghei_PBANKA_020880	FEKGNHSTNENDK-----NKDLLAIEENI	183
P.yoelii_PY01242	-----	
P.falciparum_PFA0220w	NNINMKS-----NTNNNNNNN	204
P.vivax_PVX_081545	ADGGRAQRMGEWGRSKRGAKQSAGGDPDKPHKSHKPGKMRKMSKMKKVR	297
P.knowlesi_PKH_020900	AEVKAEGMGEWGGHSGRGGKHSS-----RRGDSKSAHKPR	276
P.chabaudi_PCHAS_020720	NDHFENT-----NEQLNKNQD	168
P.berghei_PBANKA_020880	NDHFESS-----DGYENEHSN	199
P.yoelii_PY01242	-----	
P.falciparum_PFA0220w	NNNNKDDDYDDNNYHKNHSDSINNSINYNDIHRKEKKKNKNTHNEKKY	254
P.vivax_PVX_081545	KMQSMRKSRRGGENGKSLIGCRSGSVTTQSSSPDQDCYQAEERKAKKK	347
P.knowlesi_PKH_020900	KVEKMRKSRRKKNWKNCLIGCRSGSVMTQFSSPDHDCYQAEERKKEKKK	326
P.chabaudi_PCHAS_020720	KDENIGK-----PNKREDSDEHESGRKKRSVKDN--	197
P.berghei_PBANKA_020880	KSEKGDEYENEHSNKSEKGDKPENEHSNKSTEMNEHGNDGKKRSVKDN--	247
P.yoelii_PY01242	-----	
P.falciparum_PFA0220w	ISNMYNFQYDDYDKKKKKNTLETYDSDTKNSDIFINTGFLPYSLNKKKN	304
P.vivax_PVX_081545	KKKK--SDINSTDGVNPQIASLFFNLSDMRKNNLSKADGVTVKKSCRHHC	395
P.knowlesi_PKH_020900	KKKKKSDMKSTEEMTNYISSLLFNLSDMRKNNLSKADGVMVKKSCSHYC	376
P.chabaudi_PCHAS_020720	-----VYEDFIREYKNLQSLFSYNQTMIKESFNP---MPTILEDNN	236
P.berghei_PBANKA_020880	-----VYEDFIREYKNLQSLFAYNKTIKENFNP---MPTISEENKN	286
P.yoelii_PY01242	-----	
P.falciparum_PFA0220w	NKKRKGGKKNEQEENRHNVNVDNMDDDDDNNDNNDDDDSNNNNNDDDS	354
P.vivax_PVX_081545	KGERTRGADGPSRGGQLEGASPPDRKNGSPANGAARGSAKISSKYS--A	443
P.knowlesi_PKH_020900	RVERTR--DG-ADVEGQRRGDECGDECGECDGASPRHGKNGSTADGAA	423
P.chabaudi_PCHAS_020720	YSDKNGSERNEDMKIPFDENVDKNATPRNHINPDYSAAYNLDDAVISDV	286
P.berghei_PBANKA_020880	GSDIK-----	291
P.yoelii_PY01242	-----	
P.falciparum_PFA0220w	NNNNNDDDSNNNNDDDDSNNNNNDDDSNNNNDDDSNNNNDDSNNN--	401
P.vivax_PVX_081545	KNYGKNSAKICGKYSAKNCAQNSGKNIRKKIGKHIGKYIADDEDGEDGEEV	493
P.knowlesi_PKH_020900	GRSGKFSRKFCVKRSKAKNCAQNEGNKICKKISKHMNKYIAEDEDGDE---	470
P.chabaudi_PCHAS_020720	NKTSEKVSEEVHETMDNKVDETDHPHFSNQLNYPTNNDVQEDENKSL---	333
P.berghei_PBANKA_020880	-----MDKKMNKTEERNHSDFSNQLNYPTNNNLKDDENKSR---	328
P.yoelii_PY01242	-----	
P.falciparum_PFA0220w	-----NDDSNNTGSFFKNMKIQSHVINNKYDNTNDYL	433
P.vivax_PVX_081545	LSEGEILSEGEVLGEVGETQGERHSGKGGKSGKNSSGREKDTAERDD	543
P.knowlesi_PKH_020900	-----GKEEILGETQGERQKKKKKNVKNNTN-RKKDTAE---	503
P.chabaudi_PCHAS_020720	-----LRERGNENDKENGVEQTQNNDTN-----	355
P.berghei_PBANKA_020880	-----LSEQDDE-----IEEKSSDNN-----	344
P.yoelii_PY01242	-----	
P.falciparum_PFA0220w	DDLESFEYNKTKKKKKKNDTIDDVFNKRNNDNTKIKYNNNNNNNNNNV	483
P.vivax_PVX_081545	RSGNCSSREAGTKNFAQGRGRTLHDDSRAPAPEWHFVRYGRGDEASGQSA	593
P.knowlesi_PKH_020900	-----TKNYAQGRRETLHDDSRTPAPEWHFVRNRRGEEISSESA	542
P.chabaudi_PCHAS_020720	-----LGEEKEHMEELKKEIKHKEFENQE-RKQNKKEKEMNDASA	392
P.berghei_PBANKA_020880	-----FGEKGEHEELKKEIKNKESENRENKKINKQKETSDINA	382
P.yoelii_PY01242	-----	
P.falciparum_PFA0220w	LEYELNYVHNSFDTHPKKWSHSLPDNYEGEKKKKKNKTKKRHDMNFYS	533
P.vivax_PVX_081545	GGNQRGSHRGSSSGSGSNRSCDSGSHRSRDNGSSSGCKRDSRRCSHRRS	643
P.knowlesi_PKH_020900	GRNHR-----RRYRSVDSGNSSGCERGSRRR-RSSS	572
P.chabaudi_PCHAS_020720	PENE-----KNIDEISKKKENHYIQKKNT	416
P.berghei_PBANKA_020880	PESEENIIL-----EKEKSDSISKKKENLYMQKKNT	414
P.yoelii_PY01242	-----	

P.falciparum_PFA0220w	IDKNNLDEQDLFSNQEALTILKNFAKENNVSHPHYKEKKINKNKKTS----	579
P.vivax_PVX_081545	RRRSRRDGGPLFGGNVFLASETEGEGSPSISPKNRNVVNSRRALTPSG	693
P.knowlesi_PKH_020900	ISVNLPEGGPPLGGNVFSADETEGEGAASVSPKRDARVVAKLGRAEPPSD	622
P.chabaudi_PCHAS_020720	--ITQLQEKSSSSNTSIASIINDNYDDHKISKNNIEHVGTQFQFSN----	460
P.berghei_PBANKA_020880	NKLQKGKEQNSSSNISVASIINDNFDNYNISKNNISIKIDKTKTYN----	460
P.yoelii_PY01242	-----	
P.falciparum_PFA0220w	-----SNSMYKDYSNNINNNNNNDENN--YKDIVSNYDNNYEYDKKN	621
P.vivax_PVX_081545	SLIGSEVSTSRSDCADRDAPFKKNSKQSQPWGERTMNS--NSYSLHDAEG	742
P.knowlesi_PKH_020900	SHSVSEVRKAARCDSDVRDAPFKNHSKQSEFLYERTMKSSNSYSLRDGED	672
P.chabaudi_PCHAS_020720	-----DSIKKKHHPIDRNEYKSSSPKNNRTNTEIKKGITESDNLKIKK	503
P.berghei_PBANKA_020880	-----DLTKKKDGSIDNNEYKNDASKNYRPSIEINKSKNESDNIKIKK	503
P.yoelii_PY01242	-----	
P.falciparum_PFA0220w	KNKIDVEQNVHMINYINVNNKQTNINNHNDNNFNVNKELSINDNVYENFI	671
P.vivax_PVX_081545	FADAHQQRYAEEDSHPQKESEQRGKKNKKKNSKGIDLTSLPERGSIGEV	792
P.knowlesi_PKH_020900	FEEAQQERYGEEERHPQRELEQREKKNKKKNSKEIDVDGTLPERGIIAQV	722
P.chabaudi_PCHAS_020720	IEKETD----DPNLYINILDENSKDNKKKDLYLNNFYLSKQD-----	542
P.berghei_PBANKA_020880	AEKQEGNKMNDHNFNINFLDES IKDNSKTDLFLNISDSFSKHD-----	546
P.yoelii_PY01242	-----	
P.falciparum_PFA0220w	REYKNLQSLFSYNKNKIEDHFNPLTRIEKNKEDNIVLENNINKFILNAH	721
P.vivax_PVX_081545	DRRGHRQSI THEKEHKESTSKYTHVIASQKKKKNSSEIHDEGSIHDEKK	842
P.knowlesi_PKH_020900	DQGRHRPNSTYEKETKVSTSKYAHTIASPQKKKRNSEIQDDGVSNDDEKK	772
P.chabaudi_PCHAS_020720	-RSSKIQSAKNVE--EVNGKIKKVKKHQASSDESSEVNEM-----DH	581
P.berghei_PBANKA_020880	-NNDNDNKLKGVKNI EVNSQIKKTSKHRVIYDDNSEVDEMNQIDINLNDH	595
P.yoelii_PY01242	-----	
P.falciparum_PFA0220w	EGLSKKMLS YHMDEQEDVQGMKSIEDDNKGGDNKGGDNKGGDKKGGDNKG	771
P.vivax_PVX_081545	TSVKDNVYENFI REYQNLQSLFSFNKDKIGEEKGEE-----EEERK	883
P.knowlesi_PKH_020900	TSVKDNVYENFI REYQNLQSLFSFNKDKMEDDNDNDNDNDNDNGDEDDDD	822
P.chabaudi_PCHAS_020720	ENYSKINKSLTEQHRKKDTKKKKVEKKKNNNNNTENTFDSKHLNDDNFS	631
P.berghei_PBANKA_020880	ENYNKINNNMNEQNRKGD IKKKKVEKKKKNSN---KNTFISNHLNDDNIS	642
P.yoelii_PY01242	-----	
P.falciparum_PFA0220w	GDKKGDDKKGDDNKDDDDDDDDDDDDDDDDDESTISYSKSDLSKIVEY	821
P.vivax_PVX_081545	KERKKEGGEKDDRKDDPFGGSLSRALDKDVEENLREDAPSKNHLSTMLTT	933
P.knowlesi_PKH_020900	DDEEEEDKKKKKGGDSYGESRSRVLRDGNESLCEVLSKNHRSRMLAN	872
P.chabaudi_PCHAS_020720	LS-----YFTED-KKEGYDSIHEAKEVLQYVNGKVK	662
P.berghei_PBANKA_020880	LS-----YFTEENKKEDYDSIHEIEKSSNEYIHVNNK	674
P.yoelii_PY01242	-----	
P.falciparum_PFA0220w	INNDME-----EMTRFSNNNSVLQKKKQNKKKKNEQNKDNILTNSKH	865
P.vivax_PVX_081545	DGYTNSKGEAPPVDAKRGKSSDPNRSDKRVESLEELSYNF GHYFNESID	983
P.knowlesi_PKH_020900	DGYANSKGEKSPRGAERGKPSNPQSDQGVDSLEQLSYNIGQYDNTESVG	922
P.chabaudi_PCHAS_020720	KHIRENK-----ETNKYKNVDPTLVKKIKNITKHTDFNINNLDQLCKL	706
P.berghei_PBANKA_020880	KDIRDNK-----ETNKYKNIDPTIMKKIKNINKGTGLNINNRVCKLNPS	718
P.yoelii_PY01242	-----	
P.falciparum_PFA0220w	NSTHTINCKNKDLKNLSTSTNIMDES IHKNNNDNMNSNKNNDNMNS	915
P.vivax_PVX_081545	KEEIGVVPRKDSLRSRDKLGGNDDAGSSGVVPMGDNRRGKTHSSGKENA	1033
P.knowlesi_PKH_020900	NIEVGVVVRDNPSPSRGKGLCGNDDTSSGVVPIGDNQQDKSHSDGKENE	972
P.chabaudi_PCHAS_020720	-DPPIKHSKHHKETTKKKSVRQSSS-SVKKKHTKHSSKKNKKEDDPNDS	754
P.berghei_PBANKA_020880	-IKEIKQVRNHKEIAKKKTLSPNSSFIKKKKHIKHS TKKNKKEDDANDSK	767
P.yoelii_PY01242	-----	
P.falciparum_PFA0220w	NNNNNDNMNSNNNNNDNMNSNNNNNNNNYSSSDMKIIESNNIPYSPNN	965
P.vivax_PVX_081545	SGMNEAVTHEHGVEKNRRGQKSADVDPSEGVGGKKNKTEGSSRQGRRESS	1083
P.knowlesi_PKH_020900	SGTNEGVTTHAHGVGK----ESEDMDASEEGTVKKNRGTGGSTRHVRRESS	1017
P.chabaudi_PCHAS_020720	IKEWESKFIEEDE-----EEDDEY	773
P.berghei_PBANKA_020880	IKKWESKFVEENEEMEENEEMEENEEMEENEEMEENEEMEENEIEEEKEED	817
P.yoelii_PY01242	-----	

P.falciparum_PFA0220w	NMKKKKKKKKLKEKLNNNIKYNEIITK-----	994
P.vivax_PVX_081545	GGSVQRMROIGSLHSGAKGNSQDGEADMNGVRGENADLLVRDDLGSQDV	1133
P.knowlesi_PKH_020900	GGNVHRIKEAANLESSTKVKSQDEDTMNRVRSKNSDLLIWNEQVQPG--	1065
P.chabaudi_PCHAS_020720	NDPYLTIINEKQKKKESKHSKKKKIYMTN-----	802
P.berghei_PBANKA_020880	NELYWETINEKKKKK-KYSKKKKSQIMN-----	845
P.yoelii_PY01242	-----	
P.falciparum_PFA0220w	-----THIFPTNKIIQ	1005
P.vivax_PVX_081545	VGRGDIVGRGDVVGRGDIVGRSDVVRDNPLARADSLGKVSLSNKTAPN	1183
P.knowlesi_PKH_020900	-----DMLLPDLSLRKGSLSKNKTVPD	1087
P.chabaudi_PCHAS_020720	-----ESFKNSISTVD	813
P.berghei_PBANKA_020880	-----ESFNNSISSVD	856
P.yoelii_PY01242	-----	
P.falciparum_PFA0220w	HDKGVEYETTNSKHLHKNINNIYNQSEQNWSLHEDLLKEVLTKEEYNEK	1055
P.vivax_PVX_081545	ENNYFEYEANNSKLLHKNMSNLYSMNGSNLTFHQDFFKDILSLNEEKKK	1233
P.knowlesi_PKH_020900	ENNYFEYGAKNSKLLHKNMSSMYSMNGSNLSFHQDFFKDILLNDEKKK	1137
P.chabaudi_PCHAS_020720	DTQSGQEDDENDQAMQLSFFNFFDKQKNNQHIVNLQKEDNEIEKDELNYK	863
P.berghei_PBANKA_020880	DIQSEQEEDENDKAMQLSFFNFFDK-----	881
P.yoelii_PY01242	-----	
P.falciparum_PFA0220w	LIKKNKNKNSKINKKTVDNKETHLHKQIAEKYDNIHTYIVETR-----	1098
P.vivax_PVX_081545	EKKKKKSSKQHSFDQFDELSYFADVKSGKHSKHRNCTPQEDFAKGFD	1283
P.knowlesi_PKH_020900	EKKKKKSSKQHSFDNIQHLPYFEETKKGKQNKEN-----RKD-----	1174
P.chabaudi_PCHAS_020720	KKKKKKDNKKTNDILKDD-----VKFEHNINNFFSSFDWNE-----	900
P.berghei_PBANKA_020880	KKKKKKHSQKSNLILKEN-----LKLNDINNFFISFDWKN-----	918
P.yoelii_PY01242	-----	
P.falciparum_PFA0220w	KDKYSPSDHEKQNSFIKERVLSKKKIKGKNSKRNIKMVSRNKENKKER	1148
P.vivax_PVX_081545	KREKGGDRKQSSGGKSAAGGRSLRGEEENDEEGDDEEENDDGEVERGK	1333
P.knowlesi_PKH_020900	-KKKGNDDDDKHGED-----DEYNDDGDVDDVDDGDDEVHREE	1213
P.chabaudi_PCHAS_020720	NLKKKNKDDLYN-----LNNENN-KDIYSVNKIEKKINKP	934
P.berghei_PBANKA_020880	NLKKNNKEDLYN-----FNNENNTKDIHSINKIEKKMSKQ	953
P.yoelii_PY01242	-----MSKQ 4	
P.falciparum_PFA0220w	TMKSQNDNHLNHNHESDDNNSIENSYEESAMYDENNSSIHDDNSKKEKFS	1198
P.vivax_PVX_081545	AAAEEDSDDYGDCHDGHGHDYDGESEGDGD-SYGHEDLRDGDHLEG	1382
P.knowlesi_PKH_020900	AAGEDIYKE-----ESDDYGNDAEGVEGRGDDSYDQEEHLNGDHLER	1256
P.chabaudi_PCHAS_020720	YLHISKEDNIN-----EQIDRLRREIAQNNILTE	963
P.berghei_PBANKA_020880	YIHTPKEDNLN-----EQIDILRKKSQNILQN	982
P.yoelii_PY01242	YVHISKEDNLN-----DQIDVIRTNYSQNNILRN	33
	. : . . . :	
P.falciparum_PFA0220w	NEKYHERAEEEEIVSDDLYQEDDSDHSNKKIKMNMKSMTSFDKDKRRYTI	1248
P.vivax_PVX_081545	SQRP SGSHKERRKKKKSQGEPTRKSKSGKLEKVKIVE-----	1419
P.knowlesi_PKH_020900	PHGQTESHKERKKKKKLQELTKKSKSGKMEKVKIVE-----	1293
P.chabaudi_PCHAS_020720	HDKHIDERKGDIRKFLHWDDGKKGKHNGKDDKK-----	996
P.berghei_PBANKA_020880	QDKYIDE-----KKGKHSKGD-----	998
P.yoelii_PY01242	RDYIDE-----KRGRHNGKD-----	49
	. :...*	
P.falciparum_PFA0220w	QNLEEIKKKSKKSINKNENDKYGYNSDYMNDSGDFAVEKKDKKKKTQQEN	1298
P.vivax_PVX_081545	-----EKNKPHKRRSVREKRNSNRTHNNSICSLGEDHLYDDAGDFL	1460
P.knowlesi_PKH_020900	-----EKNRTHKRSTKEKRKCSRTHNNSICSLDDDDFNDDAIDYP	1334
P.chabaudi_PCHAS_020720	-----EKDKLKEKDKIYKEKMDYPNYEHNRYDEDKRKHKKNHME	1037
P.berghei_PBANKA_020880	-----KDKFREKDKMKYRDKISYINYEHNRSYIDNKHKKYKNYLE	1038
P.yoelii_PY01242	-----KDKFREKDRMKYRDKISYTN-----YIDNKHKKYKNYLE	83
	*. :. : : . . .	
P.falciparum_PFA0220w	CDNKYGNKYNKCDKDK-----DKDNYNK	1322
P.vivax_PVX_081545	YNEGRGHHHDGAIGDDPPGDNLHIDVLEEFRRDSDHKGIVGYSDEVHYQE	1510
P.knowlesi_PKH_020900	YHEKRVLLDDAIGDD---NLDIDMLDDLPGK---GILAYPDEVPNQE	1376
P.chabaudi_PCHAS_020720	NYDESYNPYSIEEYN-----NYSS	1056
P.berghei_PBANKA_020880	NYDELYN-YSIEEYN-----NNYSS	1057
P.yoelii_PY01242	NHDQVYNYPYIEEYD-----DDDNNYS	107
	: :	

P.falciparum_PFA0220w	DKFLPSDQAFHYDNRKAKKKNKEDILKDQYNDEHIKEYFYSLIEGQVSKN	1372
P.vivax_PVX_081545	EGANSEKAKRRKRSRRKKEDSRDEYMGFFSSVMDNQMYKKNLNSKDFG	1560
P.knowlesi_PKH_020900	EGPNVEKAKRRKRSRRKKEDLREEYMKEFFSSVLDNQIYKKNLNSKDFC	1426
P.chabaudi_PCHAS_020720	SENE-----LKRIYNKNEKKYMDGILLNKNYDEPINNSLKKKHHKHEDKE	1101
P.berghei_PBANKA_020880	SEHEKRENELKKIYNKNERKYLDGIIILNKNYDESINNVLKKKDHKNKEKE	1107
P.yoelii_PY01242	SEDEKREKGFKKIYNKNEKKYLDGIIILNKNYDEPINNVLKKKHNKNEKE	157
	. : : : . . : : .	
P.falciparum_PFA0220w	NKNKKKKNSQRDYSLNKSTKE-----KGVKKERL	1401
P.vivax_PVX_081545	HSFGPEGEEDADEEGEDERDDRNGRDAREKHMLRNISKVIRTEINLNTIE	1610
P.knowlesi_PKH_020900	RNFALEGEEDMADAIEGEEG-----EKYMLKNISKVIRAENIDTEIE	1468
P.chabaudi_PCHAS_020720	NNYFLENNDN-STAHNNNG-----YTTKGLG	1127
P.berghei_PBANKA_020880	NNYFFQKNNDN-STVHNNSG-----YATKSLG	1133
P.yoelii_PY01242	NNNFFQKNNDNSTVNNSG-----YATKSLG	184
	.. : : : . .	
P.falciparum_PFA0220w	LHNKHFKETDSEED-----QNNKKNNNIYLKKNYDQENEKDNEYEN	1443
P.vivax_PVX_081545	LLRNEILQTKMSKESLPVSEKKKGAHKIDNSVEATHRKKKSYDKEEKRA	1660
P.knowlesi_PKH_020900	LLRNEILQTKMSKESVSHVEKKSAHKISNPLEMTHRKKKSNDEKEEKRN	1518
P.chabaudi_PCHAS_020720	FFSSNFFHLKKLIE-----KKNVENNDSLKRENGVDKFYEKNENNAK	1169
P.berghei_PBANKA_020880	FFSSNFFHLKRLLE-----KKNVETNDPLKRENELDKFYEKNE----	1171
P.yoelii_PY01242	FFSSNFFHLKRLLE-----KKNVETNDSLKRENELDKFYEKNE----	222
	: ..: . : : : : : : : : : *	
P.falciparum_PFA0220w	EKS-----YKKSTRPYEEDHTPYRKQNIQD	1469
P.vivax_PVX_081545	SFMGGPRVDSSEDRGVDPRTAAGSHHNAIPSTYAKEEDDYAFHHEKKK	1710
P.knowlesi_PKH_020900	SFK-----LDFKTGKDGTHHNAFTSHYAKDEDDYYTIHHEKKK	1556
P.chabaudi_PCHAS_020720	NE-----KKKLSSFDILQKFINNKG-DTKY	1193
P.berghei_PBANKA_020880	-----KKKSSSFDIFQKFINNKGHDMMK	1194
P.yoelii_PY01242	-----KKKSSSFDIFQKLINNKTDMKC	245
	: : : :	
P.falciparum_PFA0220w	WSSYTKDKENKLNMDDDINMNKGNDQDVNRTYKNEKNKEEDKYGKNEKNE	1519
P.vivax_PVX_081545	RESHVKLRDVSFEEDDLRRPKGRSRIAGRDTVHFSNDEGSGIPPTHGNS	1760
P.knowlesi_PKH_020900	RESHVMRDINFEDERFKKSKGRNKLAGRDRVHSFNEEGRGIPPTHGNN	1606
P.chabaudi_PCHAS_020720	GKIDDSYKRGNNDAELYHISRRKGSINEYGDPHETAQISDGYSEKAGHN	1243
P.berghei_PBANKA_020880	GKISNFYKRESHDYSDIYYTSRGKESFKECIDPNDTEQISDAYSEYIEHF	1244
P.yoelii_PY01242	EKMSNSYKRESNDYSDMYTSREKENLKKYIYPNDTPQISDAYNDYIEHF	295
	. : : : . . : :	
P.falciparum_PFA0220w	K-----YDKYDKYKYEKYDKYKKNQHDPLYDNIKNKYDNDN	1560
P.vivax_PVX_081545	SSIGGG--LHLNKSLEFFSSNFLSLKRLIERKSGDPHYDAKREDDYVDRA	1808
P.knowlesi_PKH_020900	NSAGGGGGLHLNRSLEFFSSNFLSLKRLIERKNVDTNDAKREDDYVEKG	1656
P.chabaudi_PCHAS_020720	D-----IDEYTNFQTYRR-----NPDERDRRYKDKREYHEN--	1275
P.berghei_PBANKA_020880	D-----IDKYGNFRKYKR-----NKDDKNRKYKDRHEYHEN--	1276
P.yoelii_PY01242	G-----IDKYEKFQTYKG-----NKSDKDRTYKDRCESHENGK	329
	: . : : . : : : . : :	
P.falciparum_PFA0220w	KGLEFFSNFFHIKKFIEKKENENVHMSKIENSQKEEELNHNKRNLLNSSG	1610
P.vivax_PVX_081545	D-----EEVNFVKKKKNSYDDYEHQPQMVHKNSHKNSRGLRFDNFGIN	1853
P.knowlesi_PKH_020900	E-----EEVDIDVRRKKNSYDDYEHQPMMNKSSHKISRGLRFDNFGIN	1701
P.chabaudi_PCHAS_020720	-----EKHEKNSKHKKHAEYKWKKINKELINDHTDHIE	1308
P.berghei_PBANKA_020880	-----EKHEKNSKHKKPAEYKCKKINKELLNDRIHNIE	1309
P.yoelii_PY01242	NGKNEKNE-KNEKNEKSEKNSKNSKNKRAECKSKINKELLNDRIQNIE	378
	: . : : . . : : .	
P.falciparum_PFA0220w	K-TEKLEKFLGLYKENNEAMDFYKSVLIEENNSMNIKSKINKNNIIDDR	1659
P.vivax_PVX_081545	KRSSDSLFSFYRSFIQEG-----KSGKGD-HRKKGRGRDSEERFDSGF	1896
P.knowlesi_PKH_020900	KRNSESLNFYRSFIKEDG-----KSGKGDHHRKKGRGRDSEERFDSGF	1745
P.chabaudi_PCHAS_020720	D-DEKSVKFRKKFNYN-----LTRTGSNESMINEF	1337
P.berghei_PBANKA_020880	Y-DDKSLNFQKKYISN-----LTRTDSNETMVNEF	1338
P.yoelii_PY01242	Y-DDKSLKFQKKYISN-----LTRYSNEMLLNNF	407
	.. .* : : . . : : .	
P.falciparum_PFA0220w	MKDNIS-----KINRYNSDDTYIKVENNYDNKKEMNN	1691
P.vivax_PVX_081545	GAGFGSGFGAGFGG----GYVTGFGSDSGEHTGAARRAPRGEDAENADD	1942
P.knowlesi_PKH_020900	GTWFGTGLGSDLVGERRGAARRAQRSSEDYMEDVEEVGVSANGQDTHNVDD	1795
P.chabaudi_PCHAS_020720	GINKLN-----PQFSDEEKN-----ENNNINK	1359
P.berghei_PBANKA_020880	SMDKLN-----HQSNNEDKNENNDNFNKENNNGFNK	1370
P.yoelii_PY01242	SVDKLD-----HQFNNGDEN-----YENCENYEN	431
	.. : : .	

P.falciparum_PFA0220w	SDELNGNNNNNNNNKNNNNNNNNNNNNNNNNNNNNNNINNGGDKNRNRFNN	1741
P.vivax_PVX_081545	MEDVEDADDAEADAGVTGDGEHDDDLHADSVADDPVRSARRRKGEKAFSA	1992
P.knowlesi_PKH_020900	EGDNDNKDDDDNDRGDDNDKGDDHEAAYGEDYADDRIRSIRRRKGEKVFS	1845
P.chabaudi_PCHAS_020720	YDE-----KDD-----EIYENSENFLYS	1377
P.berghei_PBANKA_020880	YENNNEFNKYENNNGFNKYENNNEFNKYENNNDLK KYDEFYENSESFLYS	1420
P.yoelii_PY01242	YEN-----YEN-----YENN-----ENN-----ENSESFLYS	453
	: .. :.	:
P.falciparum_PFA0220w	NN-----	1743
P.vivax_PVX_081545	GRAADGVAGRAAGRSAKAARGRAKRGEQLHSGEQLHSGEPLHSEDPLHS	2042
P.knowlesi_PKH_020900	GREEDVVAGRAGSRSEKGTLRDRAKR-----	1871
P.chabaudi_PCHAS_020720	-----	
P.berghei_PBANKA_020880	-----	
P.yoelii_PY01242	-----	
P.falciparum_PFA0220w	-----	
P.vivax_PVX_081545	EDPLHSEDPLHSEDPLHSEDPLHSEDPFHSEDLHSEDHLHSEDHLHSED	2092
P.knowlesi_PKH_020900	-----DAVHG-----QSTEDLHSGEKLDAGEKLQNEE	1898
P.chabaudi_PCHAS_020720	-----	
P.berghei_PBANKA_020880	-----	
P.yoelii_PY01242	-----	
P.falciparum_PFA0220w	-----IYMCKNVKNIIISLELSNEEKINEVR	1769
P.vivax_PVX_081545	HLHSEDHLHSEEPLPSEDHLRCGP GKSHKNVKNIILSLELSNEDIHKVR	2142
P.knowlesi_PKH_020900	QSVNVEELH-----GSGRNHKNVKNIILSLELSNEDIKYVR	1935
P.chabaudi_PCHAS_020720	KILFYSSSEKKIMDEILNTLYIYPQLYVSCIICLFYLFILDNDLFERN	1403
P.berghei_PBANKA_020880	-----VDNFKIVKNIIISLELTNDNKINKIR	1403
P.yoelii_PY01242	-----VDNFKIVKNIIISLELTNDNKISKIR	1446
	*****:*****:*** **	479
P.falciparum_PFA0220w	KILFYSSSDEKKYIMNEILNILIYIPQLYVSCIISLFYLFILDNEIFEKY	1819
P.vivax_PVX_081545	KILYYSSSDEKKLMINEILNTLYVVPQLYVACIIICLFYLFILDNELFERN	2192
P.knowlesi_PKH_020900	KILYYSCSEKKLMINEILNTLYVVPQLYVACIIISLFYLFILDNELFERN	1985
P.chabaudi_PCHAS_020720	KILFYSSSEKKIMDEILNTLYIYPQLYVSCIICLFYLFILDNDLFERN	1453
P.berghei_PBANKA_020880	KILFYSPSEKKKIMDEILNTLYMYPQLYVSCIICLFYLFILDNDLFERN	1496
P.yoelii_PY01242	KILFYSPSEKKKIMDEILNTLYMYPQLYVSCIISLFYLFILDNDLFERN	529
	:** *:* **::** **:*****:***.*****:***:	
P.falciparum_PFA0220w	FNADDLVLYLFNEKIDFRYAWEFLKTYLFYKYKYSDNTHTKGSIIYIKKS	1869
P.vivax_PVX_081545	FNSDDLVLVLFNEKIDFRYAWEFLKTYIFYKFKFSSGTHSSEYKNCVRS	2242
P.knowlesi_PKH_020900	FKPDLLVLVLFNDKIDFRYAWEFLKTYIFYKFKFSSGTHSSEYKNCVRS	2035
P.chabaudi_PCHAS_020720	FSPEDLVLYLFNDKIDFRYAWEFLKTN---EQLKNEKLK---SEKN--GDINS	1503
P.berghei_PBANKA_020880	FSSENLIYLFNP KIDFRYAWEFLKTYLFYKFKFSNSNSDDGYNSYICE-	1545
P.yoelii_PY01242	FSSKDLVLYLFNP KIDFRYAWEFLKTYLFYKFKFSNYSNSDDEHNSYICE-	578
	...:**** *****:***:*. .:..	
P.falciparum_PFA0220w	PRNSIKREDSNM YADQS VMAENIKKNYLNEG NQKDD---DNKN NYDDKEN	1916
P.vivax_PVX_081545	VGSDGKQG SVDLEEVSPPGG EAGRAAG-EATDKAAGEATDKAAAGEADR	2291
P.knowlesi_PKH_020900	VGSEGKRASVDKEIGENSLEG AHAFSGGN SHRDMASADRAVDRTADK	2085
P.chabaudi_PCHAS_020720	ELFDEPLKF DAKSIDGKD AEGLKN---EQLKNEKLK---SEKN--GDINS	1545
P.berghei_PBANKA_020880	-----NDSIGKEKVA-----KIEKGR---IEKG--GDINS	1570
P.yoelii_PY01242	-----NSSVGKENVANTEKGRIEKDR IEKGR---IERG--EEANS	613
	: .. :	:
P.falciparum_PFA0220w	NYDDKENNYDDKENNYDDKNKNYDDKNNN-----	1945
P.vivax_PVX_081545	AVGEEADRD PPNVADRAPDRASDRAPDRVSDRGG-----	2325
P.knowlesi_PKH_020900	TADKTADKTDRVADRVADKTDVKGVEREGDGV TASTADGICDRAS	2135
P.chabaudi_PCHAS_020720	LISEKDDTRGNSTLDGIEGNEKTKRKNKA-----	1574
P.berghei_PBANKA_020880	LVSEKDDIEGN-----KKIQKI-----	1587
P.yoelii_PY01242	LVSEKDDIEEN-----EKQKQI-----	630
	. . : :	.
P.falciparum_PFA0220w	-----YDDKKN NYDDKKN NYDDKNKNYDYNNNNKDDDDSI NVSSSLK	1987
P.vivax_PVX_081545	--LRERHIQDKMF DVFFEGYHSSVD PADNYFCERRSHMARGER-----	2366
P.knowlesi_PKH_020900	DSARERHIKKMF DIFFEGYHSSVD TDNYFCERKRHMARGER-----	2179
P.chabaudi_PCHAS_020720	-----KNII CDNFFESYNSSGSDMDNYVYC PNKIEE--RNKN NESD	1614
P.berghei_PBANKA_020880	-----KNTIYDNFFESYNSSGSDMDNYLYY QNDKIEE--QNKINESN	1627
P.yoelii_PY01242	-----KNTIYDKFFESYSSSGSDMDNYIYKNGKNGKNCNDKNYEPN	672
*	*

P.falciparum_PFA0220w	GIHKNTFDPFPEKHSNNSLMDSGDDYLCDMNNLSNNKKDIYILWTFESS	2037
P.vivax_PVX_081545	-----DAHDLWKSIEKL	2378
P.knowlesi_PKH_020900	-----ARTGNDSSDLWKSIEKL	2196
P.chabaudi_PCHAS_020720	-----KQSERKIDEIWNKIESN	1631
P.berghei_PBANKA_020880	-----KKLERKIDKTWNKIETN	1644
P.yoelii_PY01242	-----KELEKKIDKTWNKIEIN	689
	. * . : *	
P.falciparum_PFA0220w	KCVGYNECKTLLSLCLKNENETCINNISASKLRSLVISIWSNIPSSKPKR	2087
P.vivax_PVX_081545	KCVGYKECKTLLKLCLKNEKQKNINQLKASKIRNLVISIWSNIYTSRPKK	2428
P.knowlesi_PKH_020900	KCVGYKECKTLLKLCLKNEKKNINHLKASKIRSLVISIWSNIYTSRPKK	2246
P.chabaudi_PCHAS_020720	KYVGYNECKMLLKICLTNENIQNIKMLSESKIRSLVISIWLNIRCSRPKK	1681
P.berghei_PBANKA_020880	EYVGYNECKMLLKICLSNEQVQNIKLLSESKIRSLVISIWLNIRCSRPQK	1694
P.yoelii_PY01242	EYVGYNECKMLLKICLSNEKQVQNIKLLSESKIRSLVISIWLNIRCSRPKK	739
	: ***:*** **.:**.*: * : . **:.***** ** *:::	
P.falciparum_PFA0220w	SFIKLIFNWNNKKDDLHKKKNLFYLLKSEKKNNKNLSKICFNFLNYLI	2137
P.vivax_PVX_081545	SFLKIIFNWLNCKKDDLKKNKLLYLLQAEKKKNVHLQKICFSYFLMYII	2478
P.knowlesi_PKH_020900	SFLKIIFNWLNCRKDDLRSKKNLLYLLQEEKKKNIHLQRICFTYFLNHII	2296
P.chabaudi_PCHAS_020720	YFLKLIYNWNNKNDDFNKKKNLFLNLLKAEKKHNKNITKACFNFLKFLI	1731
P.berghei_PBANKA_020880	YFLKLIYNWNNKNDDFNKKKNLFLNLLKAEKKYNKNITKACFNFLKFLI	1744
P.yoelii_PY01242	YFFKLIYNWNNKNDDFNKKKNLFLNLLKKEEKYNKNITKVCFSYFLQFLI	789
	.::***: * :*: :****: ** : *** * : : : *.*** .:*	
P.falciparum_PFA0220w	KYKDNCSNDIYYILYLIDENELKIYSKNFIQNHKINFNQFISIWNIMCIL	2187
P.vivax_PVX_081545	KHKEECSNDLIYSLYLIEEGELKKYSRDFQMNRIQVNFISIWNVMCIL	2528
P.knowlesi_PKH_020900	KHKEECSNDLIYSLYLIEEDKLKKYSKEYFEMNRIDVNFISIWNVMCIL	2346
P.chabaudi_PCHAS_020720	KYKENCSDIYYTLYLIDELKLYSKKFIQNKININQFISIWNIMCIL	1781
P.berghei_PBANKA_020880	KYKENCSDIYYTLYLIDENELKLYSKKFIQNKININQFISIWNIMCIL	1794
P.yoelii_PY01242	KYKENCSDIYYTLYLIDENELKLHKKKFIQNKININKFISIWNIMCIL	839
	.::****:* *****: ** :*: :*: :*: :*: :*: :*: :*	
P.falciparum_PFA0220w	FWDTEINNFTFLQKN-KYYYYDFMLIFLKTIFYDYINVNRDMREIMKMKL	2236
P.vivax_PVX_081545	FWDTKKINDFAFLHKS-KFYYYDFMLIFFRAFYEYACVNKGVGALYLR	2577
P.knowlesi_PKH_020900	FWDTKKINDFAFLHKS-KFYYYDFMLIFFRAFYEYVCVNKGIERGALYLR	2395
P.chabaudi_PCHAS_020720	FWDTKKIHNFTLSNKNNTNNYYDFMLIFLKTIFYEYIFT-ANSQFGQNIYL	1830
P.berghei_PBANKA_020880	FWDTKKIHNFTFFKKNNNNYYDFMLIFLKNFYEYIFP-ANSQIVQNIYL	1843
P.yoelii_PY01242	FWDTKKIHNFTFFKKNNNNYYDFMLIFLKNFYEYIFP-ANAQIFQNAYL	888
	****.:*:*: :* . : *****: ** : *	
P.falciparum_PFA0220w	KRTFLTGYHHDVEEPSQEHMSLYQE-----KNNIHNQDNRLS	2273
P.vivax_PVX_081545	RAGAAAAASAATAAT--ATAGAD-----AADG-AL	2607
P.knowlesi_PKH_020900	RVVPGDGGGTGTVISGVVTGGVNGGVN-----GAEGTDG	2428
P.chabaudi_PCHAS_020720	KRMSNTSGVSAISFPNRYN-----	1850
P.berghei_PBANKA_020880	KKTININSISTINLTNQTHN-----	1863
P.yoelii_PY01242	KKTINFNSISTTNLTNQPHNKPSSKISSKLSSKMSKMSKLSKLSSKLSSKL	938
	:	
P.falciparum_PFA0220w	FTYMKKMSLSNSSINNKKQDKHEDQNEYNLFDIENIINNFTDFVNNEI	2323
P.vivax_PVX_081545	APPRSHSLSNCSNKKDDTQDEGDGYLNLFDVENIINNFTDFVNSEL	2657
P.knowlesi_PKH_020900	VSPRRSISLSNCSNKKDDTQDEGDGYLNLFDVENIINNFTDFVNSEM	2478
P.chabaudi_PCHAS_020720	--KLSSKASSTYAESQNKGNYNDESDDFNLFVDVENIIEKFNFNDIVNAEV	1898
P.berghei_PBANKA_020880	--QFSSKLSSSYPESENKENYNDESDFNLFVDVENIIEKFNFNDISNIEV	1911
P.yoelii_PY01242	SSKMSSKLSSSYPEP---QNYNDESDFNLFVDVENIIEKFNFNDVSNIDI	985
	. * . : : . :*:***:***: *:* * . . : :	
P.falciparum_PFA0220w	SRDNYDFSFFGATNMPIPS-----	2342
P.vivax_PVX_081545	RSSSYFDAFLGGGPAARRP-----	2676
P.knowlesi_PKH_020900	RSSSYFDAFLGGGPAARRSGVRETSSVGKQKLNSALQSARDSEAGQTGE	2528
P.chabaudi_PCHAS_020720	SSKNLMH-SPRDHNHDS-----	1914
P.berghei_PBANKA_020880	NTKNFTS-QFRDNNQDF-----	1927
P.yoelii_PY01242	NTKNFTSSQSRDNNPDS-----	1002
	..	
P.falciparum_PFA0220w	-----MSNISLAGNHTTQYEKNTRHYNNSPLTHPLWRNRQEKERDLQRI	2386
P.vivax_PVX_081545	-----EPFPADKPGEPASQGG	2692
P.knowlesi_PKH_020900	PYRGVVTGSSHRPCSTDGVESPTLQTAHYGTHFAEPFPADKAAEPSTQGG	2578
P.chabaudi_PCHAS_020720	-----HGTHPSESIESAENREE----IEQAKENTIGIKNMLKFI	1949
P.berghei_PBANKA_020880	-----IGTNHSETMESETREVEQSVHAKENTENTNGIKNMLKFI	1966
P.yoelii_PY01242	-----LGTNHSETIESE-----NANGIKNMLKFI	1026
	: :	

P.falciparum_PFA0220w	KDEEERLKRGGVPLTEAYDIENLIFLGICIKIVICRISNLLNAKSCLEQ	2436
P.vivax_PVX_081545	PQPAFAPSAQRSAYPNDSSEMNLLLLLGICIKIVICRISNLLNAKSCLEQ	2742
P.knowlesi_PKH_020900	TQSTFSPITQRSVFPDSEMNLLLLLGICIKIVICRISNLLNAKSCLEQ	2628
P.chabaudi_PCHAS_020720	IQKSK--KEKNMICCS DANELKLLVFLGICIKIVICRISNLLNAKACLEQ	1997
P.berghei_PBANKA_020880	IQKIK--KEKNIISCVDINEIKMLIFLGICIKIVICRISNLLNSKTCLEQ	2014
P.yoelii_PY01242	IQKIK--KEKNIISCVDINEIIMLVFLGICIKIVICRISNLLNSKTCLEQ	1074
	: : : : : *:*****:~::~	
P.falciparum_PFA0220w	FHYFLNHKRLGLKIFKYSHIILVYFIPFFKKYYFLWKFIEHEIDKDIMNL	2486
P.vivax_PVX_081545	FNYFLNNQKLGFRTEFKNAHMLVYFIPFFKKYYFLWKFIEHEIDVEIMTL	2792
P.knowlesi_PKH_020900	FNYFLNNQKLGFRTEFKNAHMLVYFIPFFKKYYFLWKFIEHEIDVEIMSL	2678
P.chabaudi_PCHAS_020720	FYYFINHKILGLKIFRHSIVLVYFIPFFKKYYFLWKFIEHEIDPEI IKL	2047
P.berghei_PBANKA_020880	FYYFINHKILGLRIFKHSIILIYFIPFFKKYYFLWKFIEHEIDPEI IKL	2064
P.yoelii_PY01242	FYYFINHKILGLRKFKYSHIVLIYFIPFFKKYYFLWKFIEHEIDPEI IKL	1124
	* ~::~ ~::~ ~::~ ~::~ ~::~ ~::~ ~::~ ~::~ ~::~ ~::~ ~::~	
P.falciparum_PFA0220w	IKYIMDHLENMQVENIPLSLCNINNTSNQMLPGVSNQNMNERTYAENLHN	2536
P.vivax_PVX_081545	LQDILNDLEHIQAT-----SSS-----	2809
P.knowlesi_PKH_020900	LQNIILNDLEQLQAS-----SSVGRMYL	2700
P.chabaudi_PCHAS_020720	INAIINNLENLQNN-----KNINNIGSNPMLHP	2075
P.berghei_PBANKA_020880	INAIINNLESQIN-----KNINNIENSPKFHP	2092
P.yoelii_PY01242	INAIINNLENLQIN-----KNINNIENNSKLHP	1152
	:: ~::~ ~::~ ~::~	
P.falciparum_PFA0220w	MNNIHNKFCPSSYRHTQNILNMNSTHNNSSVNNNFNKMNHSISEKMGKN	2586
P.vivax_PVX_081545	-----	
P.knowlesi_PKH_020900	GVGGTHQVCGVSNPLHVREMSQSRNVCDRLPQENLQERFTSPFGERPFVR	2750
P.chabaudi_PCHAS_020720	YTRSNTSLPEVDHFSFPSKDKNLNHTFNNFSTQYNPHDYKTSFSHSISIE	2125
P.berghei_PBANKA_020880	YIRSNTSLPKVDHLFFPSKDKNLNIFSNFNMQSKPHNYKTSFSHSISIE	2142
P.yoelii_PY01242	YIRSNTSLPEIDHLSFPLKNNKLNRRFTNLNIQSNPHYKTNFSHTPSIE	1202
P.falciparum_PFA0220w	KNDNIFSFLKSTKNNMSFDQNGRLVNSNINYMKNKNLLLCKEEQEKHTSF	2636
P.vivax_PVX_081545	-----	
P.knowlesi_PKH_020900	SSTCNSQNRGAFFTGGKK-----ISSLEHFYANGKFN	2783
P.chabaudi_PCHAS_020720	NNNSKQNFKKHKQNDLSLD-----SDLPTPDLFNPKNKKK	2160
P.berghei_PBANKA_020880	SNNNNNN--FKQNNISLS-----FDMLISDSFNVKNKKK	2174
P.yoelii_PY01242	SNNNNNNNN--LKQNNISFS-----SDMLISDPFNVKNKKK	1235
P.falciparum_PFA0220w	QSLNCNRTKNNSIQERVVYGKEINNNHNLKDINVFKYKKHEHKHGEFFNL	2686
P.vivax_PVX_081545	-----ASTDASHPPSE	2820
P.knowlesi_PKH_020900	LY-----SNEIRRMGRSGPNGEGAEEDDVYFASG	2812
P.chabaudi_PCHAS_020720	KK-----KNSEKFEDSH--KNKQQHPHDYSEFDSH	2188
P.berghei_PBANKA_020880	-----NSEKFGDTHSNENQRYFHDYSELGCS	2200
P.yoelii_PY01242	XKIKKKNSSENSKNSSENSKNSSENSEKIGDTHSNVTHGYSNGYFELGCQ	1285
P.falciparum_PFA0220w	NNMKYPLYGKNKNIMDDDNLGNNIFHPKKKNKDEFIGSFKNSSYVINDE	2736
P.vivax_PVX_081545	RDPT-----	2824
P.knowlesi_PKH_020900	RGPSPMG-----AMNTVMSVAPLKGALCRSEED	2841
P.chabaudi_PCHAS_020720	NN-----FHKNIIINNININIFRPSLSLFTN-E	2215
P.berghei_PBANKA_020880	NK-----SFHANLTNNIN---YLQPNISIFKLN-E	2225
P.yoelii_PY01242	NKNVH-----KNVHKNVNNIN---YLQNSLFLKLN-E	1314
	.	
P.falciparum_PFA0220w	FIGSFKNSSYVINDEDEHYISYDDMF	2748
P.vivax_PVX_081540	-----EEDTDENCITYDDMF	26
P.knowlesi_PKH_020900	-----EEDTDENCITYDDMF	2853
P.chabaudi_PCHAS_020720	---NYFRPSSSLFNTNENC DHGITYDDMF	2227
P.berghei_PBANKA_020880	---YLQPNISIFKLNKCDQVITYDDMF	2237
P.yoelii_PY01242	---YLQNSLFLKLNENC DQGITDDMF	1326
	:: :: *:*****	
P.falciparum_PFA0220w	RNYDSDDDS-----ISNSKNTSENFVKDFITNLHFANLDD	2785
P.vivax_PVX_081540	RNYDSENN-----EGRGSPEENCNVKDFITSLHFN---	57
P.knowlesi_PKH_020900	RNYDSEDNS-----EGRDCLDESCNVKDYITSLHFN---	2884
P.chabaudi_PCHAS_020720	RDNDNSSDD-----NYFDKGKNKICNVKEYITNLHFNN---	2260
P.berghei_PBANKA_020880	RDDDDDYDYDHDYEN-----NYFDKNKSKIHNVKEYITNLHFNN---	2276
P.yoelii_PY01242	RNEDENDEENENENENENDENGNYFDKNKKNKTRNVKEYITNLHFNN---	1373
	*: ~::~ ~::~ ~::~ ~::~ ~::~ ~::~ ~::~ ~::~ ~::~ ~::~ ~::~	

P.falciparum_PFA0220w	DNNIISKFFSTSKKLLNDQKGQEKGQNGEQK---CEQKYEQKYEHQGSSS	2832
P.vivax_PVX_081540	--NPLEERILSRRKASASQCGEKKAGRDNDDNDGAEATHGDNGEE-ETS	104
P.knowlesi_PKH_020900	--NPLEEKILSKRKGAKSVSGEKKKGKVNNDGE---EVAKNGDNVEEKNS	2929
P.chabaudi_PCHAS_020720	----LPDYPTSLKNSDENKKSEDKKKKKIK--NENQNSKEEQDNGEE	2303
P.berghei_PBANKA_020880	----LPNYLTSLKNNENDNKIEDKKKKKKKLK--DHDQSQNKNKSKEQ	2319
P.yoelii_PY01242	----LPDYLTCLKSGNANK-----KKKKKKIK--DQYQSQSQSQSQQS	1411
	: : : :	
P.falciparum_PFA0220w	VKIQNNKIINKMKYPDPLSSSESSSNYNEDKNIMYMYPNEPNYKDSKKVL	2882
P.vivax_PVX_081540	PNKEAGATVGEANGSACHQESKEDLFGSNPTRVGVFP-----ME	143
P.knowlesi_PKH_020900	YIKEGGTIMGEGNSTCPHESKEDIFGSKPTRVGIFP-----ME	2968
P.chabaudi_PCHAS_020720	EHNSNK-----TQHKENNISVKNEQNNKILK-----E	2337
P.berghei_PBANKA_020880	ENGKEENRNRSKTQLKEYNNISKLIGKNENIPFYNI PDN-----E	235
P.yoelii_PY01242	KNSKEQ----ENEEGENKNTSKLIGQNE DIPFYYISDN-----E	1446
	. : *	:
P.falciparum_PFA0220w	QKKKKKKSTINNFRINSNGPHTNEEFIEKDQSTSIIGSLGQDDSFDKIS	2932
P.vivax_PVX_081540	SSTPKQSLGQNYHFSTRAYKNGGT-----SASAGAVGVGTLTAMMM	185
P.knowlesi_PKH_020900	SSTVRQNSSGGQNYHMNSKDYKSGGAKDLPGYKKS SGGGAPGANSTSM	3018
P.chabaudi_PCHAS_020720	NAENKISKNSNTTRKYMIDENLN EYK IENKKDVSNYNNKKE SPTFEEDK	2387
P.berghei_PBANKA_020880	NSENKISKIGNSDRKYLFDNHDFEYKGNKKCLKNNYNHKEGSANFEEYK	2408
P.yoelii_PY01242	NSGNKISKTGPNARKYLLN-----NFEEK	1471
	. : :	
P.falciparum_PFA0220w	HKNTHFDHHKNPSDLTNNHMMKNVKHMKNIKQHCSN---DDYSTSKYEE	2979
P.vivax_PVX_081540	PP-----DPFANENSVMKKFPSLGVLSEDEK---VEKA EKRAEN	222
P.knowlesi_PKH_020900	PP-----DPFAENNVMKKFPSGLGPESDDEK---GENGE---VN	3052
P.chabaudi_PCHAS_020720	YR-----HNRRSSTPQNNGIKKFPTVS YEDADNSNMVDKKS QKNYIN	2430
P.berghei_PBANKA_020880	YK-----ENSKSNIHQNGINGKFFTISYENV DNLNHI IDKKTQKEYLN	2451
P.yoelii_PY01242	YK-----QNNKLN IHKNGGISKFFTISYEDVENSNM MIDTKTQNNYLD	1514
	: * : : :	
P.falciparum_PFA0220w	LVNEHTIRKNTNRRLSYAYPTQNRISDQ-----MENQKIRKNTS	3019
P.vivax_PVX_081540	PEATKGLPPTNDNR-----RAKKDSL-----KSNATAKR DGS	256
P.knowlesi_PKH_020900	FKSGKALPLNTENRK-----KEKKNDSL-----KNNTTGKGDGS	3086
P.chabaudi_PCHAS_020720	LELDREKKENFGSSKKLTNIESAKGMND-----DYNYSNNEI LNNR	2472
P.berghei_PBANKA_020880	LEMNRKKKESFSSSKRLTNIESVKDLNSG-----ENNNEI INDR	2490
P.yoelii_PY01242	LEMNRKKKESFSSSKRLTNIESVKGFNSDNGGDNDDNDENDGEAINSR	1564
	: . . : .. :	
P.falciparum_PFA0220w	LEKNVHHMNNDYDEINFTEKYFEQEY GSDQHDQRNNSMDAVNSVNVHVRM	3069
P.vivax_PVX_081540	DPNVVP SQTEQTGNHSMNGPLNEFFPYDNP MQTTTTPGSTPLSR SICIKN	306
P.knowlesi_PKH_020900	NLNI VGHQTEQINGHTINGQLNEFFPYETPMQNTTPG--ALNKNICIKN	3134
P.chabaudi_PCHAS_020720	RTIQTNIGHTVNYNRNNNNMRPDEYEN FVKD--KKNDLDIIQRKG ISLVNA	2520
P.berghei_PBANKA_020880	KSIQTNVHITNYIRNHN NIKSNEYEQFLREKKNKKLEISQRNGISIVNS	2540
P.yoelii_PY01242	KRGQASEHIGNYIRNHN NIKSKEQFLREQNKKKELEISQKKGIYIVNS	1614
	. : . . *	:
P.falciparum_PFA0220w	DGVNVHVNRMDGVNVHVNRMDGVNVVRNMHANRVSRM NHANRVSRM NHANR	3119
P.vivax_PVX_081540	PLNDNYQEVDFTTSALRRK----TKSGEGNKLRRGAAP-----	341
P.knowlesi_PKH_020900	ALNDNYQEVDFTTSSLRTK-----NRPEGNKTKKG VAP-----	3169
P.chabaudi_PCHAS_020720	TTNNNYEEINI VNNPIEKN-----KHVSSDKLIQKRP-----	2552
P.berghei_PBANKA_020880	TTNNNYEEINI INSHIEKK-----KQDSSDKLIQKRP-----	2572
P.yoelii_PY01242	TTNNNYEEINI INNTEKK-----KQINSDLLVQKRP-----	1646
	: : :	
P.falciparum_PFA0220w	VSRMNHANRVSPNNIEDIRMGGVKIKKYLMLPINKFTFENMSKR NYHPHP	3169
P.vivax_PVX_081540	-----KYLMLKVDTFELNKMTSKSLKHPP	365
P.knowlesi_PKH_020900	-----KYLMLKVDTCELKRMTRSKLHPP	3193
P.chabaudi_PCHAS_020720	-----KYLMLPIDTTTELKKMQGKGLRHPP	2576
P.berghei_PBANKA_020880	-----KYLMLPIDMIELKKMQGKGLRHPP	2596
P.yoelii_PY01242	-----KYLMLPINMIELKKMQGKGLRHPP	1670
	***** : : : * : **	
P.falciparum_PFA0220w	VGLMNLGNTCYLNSLLQALYSTVSFIVNLFLFKINETNNKV R T V P N Y E I Y	3219
P.vivax_PVX_081540	VGLVNLGNTCYLNSLLQALYSTVSFVVNLYLFNIREEPKEGGGPCSNPF	415
P.knowlesi_PKH_020900	VGLVNLGNTCYLNSLLQALYSTVSFVVNLYLFNIRQ NVKEGG-EESSNPF	3242
P.chabaudi_PCHAS_020720	VGLINLGNTCYLNSLLQALYSTVSFVVNLYIFNIDDNKELKHINNKNISN	2626
P.berghei_PBANKA_020880	VGLINLGNTCYLNSLLQALYSTVSFVVNLYIFNIDDNEKIMN LN HQSSN	2646
P.yoelii_PY01242	VGLMNLGNTCYLNSLLQALYSTVSFVVNLYIFNIDDNEKIMY SNHQVNSN	1720
	:**:*****:**:	
P.falciparum_PFA0220w	KSQM HQEN-----	3227
P.vivax_PVX_081540	AKESPASSPOGGAAOQG----GSKMNTLS-KONTLSRON TLSRLNTLNG	460

P.knowlesi_PKH_020900	VEKQKHTLGPNGTNTLKGPVHRLSNVDGLNNIMNSIGKMNTLTRNLTLNG	3292
P.chabaudi_PCHAS_020720	EMPIKNKLSFNLNN-----TN	2642
P.berghei_PBANKA_020880	IISIKNKLFFNSNNLNNP-----NN	2666
P.yoelii_PY01242	KMSIKNKLFFSSN-----	1733
P.falciparum_PFA0220w	-----TNSLDYFLEEIKSFFKNMLTTDKSYISADRVNLMLPVELNNR	3270
P.vivax_PVX_081540	TTPTNQVATAESERFLQELKNLYIIMTTTCKTYVSPDCILGLLPHELNNR	510
P.knowlesi_PKH_020900	IPPMNHIGTVESERFLQELKNLYIIMTTTNKMYVSPDSILGLLPYELNNR	3342
P.chabaudi_PCHAS_020720	MNNNNNNANLLSKRFLYELKILFKLMTTNNKKYVSPDNILGILPQELNNR	2692
P.berghei_PBANKA_020880	PNINDSYTNLLSKRFLYELKILFKLMSTTNKKYVSPYSILSILPQELNNR	2716
P.yoelii_PY01242	--INDSHKNLLSKRFLYELKILFKLMSKTNKKYISPDNILNILPQELNNK	1781
	. . ** *: * :: * .*. * *: . :*: ** ***:	
P.falciparum_PFA0220w	NQQDTEVFRYIFDKLGGSEKEFLRLIFSGVVIQKQCQKCLFISKKEEI	3320
P.vivax_PVX_081540	NQQDTELLRYIFDKLGGSSREFLRLIFSGVLVQKQCQKCFVSKKEEV	560
P.knowlesi_PKH_020900	NQQDTELLRYIFDKLGGSSREFLRLIFSGVLVQKQCQKCFVSKKEEV	3392
P.chabaudi_PCHAS_020720	NQQDTELFRTTFFQLGGSEKKFLRLIFSGVVIQKQCQKCFVSKKEEI	2742
P.berghei_PBANKA_020880	NQQDTELFRTTFFQLGGSEKKFLRLIFSGVVIQKQCQKCFVSKKEEI	2766
P.yoelii_PY01242	NQQDTELFRTTFFQLGGSEKKFLRLIFSGVVIQKQCQKCFVSKKEEI	1831
	*****: ** *: ***: .: *****: *****: *: ***: *	
P.falciparum_PFA0220w	IHDLSFPVPDISTNEKLSIQRFDDTFIQKEKIYGNKKYKCSRCKNKRNNALK	3370
P.vivax_PVX_081540	IHDLSFPVPKASERLSIQRFDDTFVQKEKIHGSKNYRCSKCNKRNNALK	610
P.knowlesi_PKH_020900	IHDLSFPVPKASEKLSIQRYFDSFVQKEKIHGSKNYRCSKCNKRNNALK	3442
P.chabaudi_PCHAS_020720	IHDLSFHVPKSSKQSIQRFDDTYIQKEKIYGNKKYKCSRCKNKRNNALK	2792
P.berghei_PBANKA_020880	IHDLSFHVPKSTKKESIQRFDDTYIQKEKIYGNKKYKCSRCKNKRNNALK	2816
P.yoelii_PY01242	IHDLSFHVPKSTKKESIQRFDDTYIQKEKIYGNKKYKCSRCKNKRNNALK	1881
	***** : * :.: ***: ***: *****: *. ***: ***: ***: *****	
P.falciparum_PFA0220w	WNEIISPPCHLILILNRYNWSFSSNEKKKIKTHVKINSKIVNNFDYKLY	3420
P.vivax_PVX_081540	WNEIISPPCHLILILNRYNWSFSSNEKKKIKTHVKINSKIVNNFDYKLY	660
P.knowlesi_PKH_020900	WNEIISPPCHLILILNRYNWSFSSNEKKKIKTHVKINSKIVNNFDYKLY	3492
P.chabaudi_PCHAS_020720	WNEIISPPCHLILILNRYNWSFSSNEKKKIKTHVKINSKIVNNFDYRLY	2842
P.berghei_PBANKA_020880	WNEIISPPCHLILILNRYNWSFSSNEKKKIKTHVKINSKIVNNFDYKLY	2866
P.yoelii_PY01242	WNEIISPPCHLILILNRYNWSFSSNEKKKIKTHVKINSKIVNNFDYRLY	1931
	*****: *****: *****: *. *****: ***: **	
P.falciparum_PFA0220w	GAIHGGISASSGHYYFIGKKSERQNKKS-SWYQMNDSVVTKANSKMIN	3469
P.vivax_PVX_081540	GGIHHGTSASSGHYYFIGKKSERDPQDEANDWYQMDDSVVSKTSSKSIN	710
P.knowlesi_PKH_020900	GAIHGGSSASSGHYYFIGKKSERQDEDETNDWYQMNDSVITKVSSKLIN	3542
P.chabaudi_PCHAS_020720	GGIHHSGVSASSGHYYFIGKKSEK-GDNSKNEWYQMDDSAITKVSSKSIN	2891
P.berghei_PBANKA_020880	GGIHHSGVSASSGHYYFIGKKSEK-CDNSKNEWYQMDDSVITKVSSKSIN	2915
P.yoelii_PY01242	GGIHHSGVSASSGHYYFIGKKSEK-CDNFKNNEWYQMDDSVITKVSSKSIN	1980
	*. ***. * *****: .. . *****: ***: ***: ***. ** *	
P.falciparum_PFA0220w	KISKDLSNDHTPYVLFYRCKQAPISPDLYF-----	3499
P.vivax_PVX_081540	RISKDPSNDHTPYVLFYRCKQAPLSPALHF-----	740
P.knowlesi_PKH_020900	KISKDPSNDHTPYVLFYRCKQAPPSPALHF-----	3572
P.chabaudi_PCHAS_020720	RISKDPSNDHTPYVLFYRCKQAPDPSLYF-----	2921
P.berghei_PBANKA_020880	RISKDLSNDHTPYVLFYRCKQAPVSPSLYF-----	2945
P.yoelii_PY01242	RISKDLSNDHTPYVLFYRCKQAPVAASFAEVAVSFVKVLTNFISSIFLNP	2030
	: *** ***** :. :	
P.falciparum_PFA0220w	---	
P.vivax_PVX_081540	---	
P.knowlesi_PKH_020900	---	
P.chabaudi_PCHAS_020720	---	
P.berghei_PBANKA_020880	---	
P.yoelii_PY01242	NPF 2033	

Appendix E – Talks, posters and publications arising from this thesis

Oral presentations:

Fitness cost of chloroquine resistance in *Plasmodium chabaudi*, Katarzyna Modrzynska, British Society for Parasitology Spring/Malaria meeting, Edinburgh (UK), 2009

Genetics of chloroquine resistance in rodent malaria parasite *Plasmodium chabaudi*, Katarzyna Modrzynska, British Society for Parasitology Spring/Malaria meeting, Newcastle (UK), 2008

Posters:

Genetics of chloroquine resistance in *Plasmodium chabaudi*. Katarzyna Modrzynska, British Society for Parasitology Spring/Malaria meeting, Edinburgh (UK), 2009

Identifying the loci of chloroquine (CQ) resistance in rodent parasite *Plasmodium chabaudi*, Katarzyna Modrzynska, Molecular Approaches to Malaria, Lorne (Australia), 2008

Publications

Paul Hunt, Axel Martinelli, Katarzyna Modrzynska, Sofia Borges, Alison Creasey, Louise Rodrigues, Dario Beraldi, Laurence Loewe, Richard Fawcett, Sujai Kumar, Marian Thomson, Urmi Trivedi, Thomas Otto, Arnab Pain, Mark Blaxter, Pedro Cravo *Experimental evolution, genetic analysis and genome re-sequencing reveals the mutation conferring artemisinin resistance in an isogenic lineage of malaria parasites* –**In press (BMC Genomics)**.

Sofia Borges, Paul Hunt, Alison Creasey, Kasia Modrzynska, Richard Fawcett, Louise Rodrigues, Urmi Trivedi, Axel Martinelli and Pedro Cravo *Genome-wide scans of selection - malaria parasites with *mdr1* duplication are selected by mefloquine, lumefantrine and artemisinin* – **submitted (PLOS One)**

Katarzyna Modrzynska, Alison Creasey, Laurence Loewe, Timothee Cezard, Axel Martinelli, Sofia Borges, Pedro Cravo, Mark Blaxter, Richard Carter, Paul Hunt Quantitative genome-wide Solexa re-sequencing reveals multiple mutations conferring complex chloroquine resistance phenotype in malaria parasite. – **ready to be submitted (Nature Genetics)**

Katarzyna Modrzynska, Natalia Bednarska, Sabrina Reignoux Paul Hunt Fitness cost of chloroquine resistance in the rodent malaria parasite Plasmodium chabaudi – **in preparation**

Bibliography

- Abdel-Muhsin, A.M., Mackinnon, M.J., Ali, E., Nassir el, K.A., Suleiman, S., Ahmed, S., Walliker, D., Babiker, H.A., 2004. Evolution of drug-resistance genes in *Plasmodium falciparum* in an area of seasonal malaria transmission in Eastern Sudan. *J Infect Dis* 189, 1239-1244.
- Adagu, I.S., Warhurst, D.C., 2001. *Plasmodium falciparum*: linkage disequilibrium between loci in chromosomes 7 and 5 and chloroquine selective pressure in Northern Nigeria. *Parasitology* 123, 219-224.
- Adjuik, M., Babiker, A., Garner, P., Olliaro, P., Taylor, W., White, N., 2004. Artesunate combinations for treatment of malaria: meta-analysis. *Lancet* 363, 9-17.
- Afonso, A., Hunt, P., Cheesman, S., Alves, A.C., Cunha, C.V., do Rosario, V., Cravo, P., 2006. Malaria parasites can develop stable resistance to artemisinin but lack mutations in candidate genes *atp6* (encoding the sarcoplasmic and endoplasmic reticulum Ca^{2+} ATPase), *tctp*, *mdr1*, and *cg10*. *Antimicrob Agents Chemother* 50, 480-489.
- Afonso, A., Neto, Z., Castro, H., Lopes, D., Alves, A.C., Tomas, A.M., Rosario, V.D., 2010. *Plasmodium chabaudi chabaudi* malaria parasites can develop stable resistance to atovaquone with a mutation in the cytochrome b gene. *Malar J* 9, 135.
- Alker, A.P., Lim, P., Sem, R., Shah, N.K., Yi, P., Bouth, D.M., Tsuyuoka, R., Maguire, J.D., Fandeur, T., Arie, F., Wongsrichanalai, C., Meshnick, S.R., 2007. *Pfmdr1* and in vivo resistance to artesunate-mefloquine in *falciparum* malaria on the Cambodian-Thai border. *Am J Trop Med Hyg* 76, 641-647.
- Amexo, M., Tolhurst, R., Barnish, G., Bates, I., 2004. Malaria misdiagnosis: effects on the poor and vulnerable. *Lancet* 364, 1896-1898.
- Anderson, T.J., 2004. Mapping drug resistance genes in *Plasmodium falciparum* by genome-wide association. *Curr Drug Targets Infect Disord* 4, 65-78.
- Anderson, T.J., Nair, S., Nkhoma, S., Williams, J.T., Imwong, M., Yi, P., Socheat, D., Das, D., Chotivanich, K., Day, N.P., White, N.J., Dondorp, A.M., 2010. High heritability of malaria parasite clearance rate indicates a genetic basis for artemisinin resistance in western Cambodia. *J Infect Dis* 201, 1326-1330.
- Babiker, H.A., Pringle, S.J., Abdel-Muhsin, A., Mackinnon, M., Hunt, P., Walliker, D., 2001. High-level chloroquine resistance in Sudanese isolates of *Plasmodium falciparum* is associated with mutations in the chloroquine resistance transporter gene *pfcr1* and the multidrug resistance Gene *pfmdr1*. *J Infect Dis* 183, 1535-1538.
- Babiker, H.A., Satti, G., Ferguson, H., Bayoumi, R., Walliker, D., 2005. Drug resistant *Plasmodium falciparum* in an area of seasonal transmission. *Acta Trop* 94, 260-268.
- Babiker, H.A., Hastings, I.M., Swedberg, G., 2009. Impaired fitness of drug-resistant malaria parasites: evidence and implication on drug-deployment policies. *Expert Rev Anti Infect Ther* 7, 581-593.
- Barnadas, C., Ratsimbao, A., Tichit, M., Bouchier, C., Jahevitra, M., Picot, S., Menard, D., 2008. *Plasmodium vivax* resistance to chloroquine in Madagascar: clinical efficacy and polymorphisms in *pvmr1* and *pvcr1* genes. *Antimicrob Agents Chemother* 52, 4233-4240.

- Barnes, D.A., Foote, S.J., Galatis, D., Kemp, D.J., Cowman, A.F., 1992. Selection for high-level chloroquine resistance results in deamplification of the *pfmdr1* gene and increased sensitivity to mefloquine in *Plasmodium falciparum*. *Embo J* 11, 3067-3075.
- Batty, K.T., Gibbons, P.L., Davis, T.M., Ilett, K.F., 2008. Pharmacokinetics of dihydroartemisinin in a murine malaria model. *Am J Trop Med Hyg* 78, 641-642.
- Beez, D., Sanchez, C.P., Stein, W.D., Lanzer, M., 2010. Genetic predisposition favors the acquisition of stable artemisinin resistance in malaria parasites. *Antimicrob Agents Chemother*.
- Beier, J.C., Keating, J., Githure, J.I., Macdonald, M.B., Impoinvil, D.E., Novak, R.J., 2008. Integrated vector management for malaria control. *Malar J* 7 Suppl 1, S4.
- Bell, A.S., de Roode, J.C., Sim, D., Read, A.F., 2006. Within-host competition in genetically diverse malaria infections: parasite virulence and competitive success. *Evolution* 60, 1358-1371.
- Belnoue, E., Costa, F.T., Frankenberg, T., Vigario, A.M., Voza, T., Leroy, N., Rodrigues, M.M., Landau, I., Snounou, G., Renia, L., 2004. Protective T cell immunity against malaria liver stage after vaccination with live sporozoites under chloroquine treatment. *J Immunol* 172, 2487-2495.
- Bennett, T.N., Kosar, A.D., Ursos, L.M., Dzekunov, S., Singh Sidhu, A.B., Fidock, D.A., Roepe, P.D., 2004. Drug resistance-associated *pfCRT* mutations confer decreased *Plasmodium falciparum* digestive vacuolar pH. *Mol Biochem Parasitol* 133, 99-114.
- Besier, S., Zander, J., Kahl, B.C., Kraiczy, P., Brade, V., Wichelhaus, T.A., 2008. The thymidine-dependent small-colony-variant phenotype is associated with hypermutability and antibiotic resistance in clinical *Staphylococcus aureus* isolates. *Antimicrob Agents Chemother* 52, 2183-2189.
- Blanford, S., Chan, B.H., Jenkins, N., Sim, D., Turner, R.J., Read, A.F., Thomas, M.B., 2005. Fungal pathogen reduces potential for malaria transmission. *Science* 308, 1638-1641.
- Boete, C., 2005. Malaria parasites in mosquitoes: laboratory models, evolutionary temptation and the real world. *Trends Parasitol* 21, 445-447.
- Borman, A.M., Paulous, S., Clavel, F., 1996. Resistance of human immunodeficiency virus type 1 to protease inhibitors: selection of resistance mutations in the presence and absence of the drug. *J Gen Virol* 77 (Pt 3), 419-426.
- Bray, P.G., Mungthin, M., Hastings, I.M., Biagini, G.A., Saidu, D.K., Lakshmanan, V., Johnson, D.J., Hughes, R.H., Stocks, P.A., O'Neill, P.M., Fidock, D.A., Warhurst, D.C., Ward, S.A., 2006. *PfCRT* and the trans-vacuolar proton electrochemical gradient: regulating the access of chloroquine to ferriprotoporphyrin IX. *Mol Microbiol* 62, 238-251.
- Burgos, M., DeRiemer, K., Small, P.M., Hopewell, P.C., Daley, C.L., 2003. Effect of drug resistance on the generation of secondary cases of tuberculosis. *J Infect Dis* 188, 1878-1884.
- Cabrera, M., Natarajan, J., Paguio, M.F., Wolf, C., Urbach, J.S., Roepe, P.D., 2009a. Chloroquine transport in *Plasmodium falciparum*. 1. Influx and efflux

- kinetics for live trophozoite parasites using a novel fluorescent chloroquine probe. *Biochemistry* 48, 9471-9481.
- Cabrera, M., Paguio, M.F., Xie, C., Roepe, P.D., 2009b. Reduced digestive vacuolar accumulation of chloroquine is not linked to resistance to chloroquine toxicity. *Biochemistry* 48, 11152-11154.
- Cambie, G., Verdier, F., Gaudebout, C., Clavier, F., Ginsburg, H., 1994. The pharmacokinetics of chloroquine in healthy and *Plasmodium chabaudi*-infected mice: implications for chronotherapy. *Parasite* 1, 219-226.
- Campbell, C.C., Chin, W., Collins, W.E., Teutsch, S.M., Moss, D.M., 1979. Chloroquine-resistant *Plasmodium falciparum* from East Africa: cultivation and drug sensitivity of the Tanzanian I/CDC strain from an American tourist. *Lancet* 2, 1151-1154.
- Carlton, J., Mackinnon, M., Walliker, D., 1998. A chloroquine resistance locus in the rodent malaria parasite *Plasmodium chabaudi*. *Mol Biochem Parasitol* 93, 57-72.
- Carlton, J.M., Hayton, K., Cravo, P.V., Walliker, D., 2001. Of mice and malaria mutants: unravelling the genetics of drug resistance using rodent malaria models. *Trends Parasitol* 17, 236-242.
- Carlton, J.M., Angiuoli, S.V., Suh, B.B., Kooij, T.W., Perte, M., Silva, J.C., Ermolaeva, M.D., Allen, J.E., Selengut, J.D., Koo, H.L., Peterson, J.D., Pop, M., Kosack, D.S., Shumway, M.F., Bidwell, S.L., Shallom, S.J., van Aken, S.E., Riedmuller, S.B., Feldblyum, T.V., Cho, J.K., Quackenbush, J., Sedegah, M., Shoaibi, A., Cummings, L.M., Florens, L., Yates, J.R., Raine, J.D., Sinden, R.E., Harris, M.A., Cunningham, D.A., Preiser, P.R., Bergman, L.W., Vaidya, A.B., van Lin, L.H., Janse, C.J., Waters, A.P., Smith, H.O., White, O.R., Salzberg, S.L., Venter, J.C., Fraser, C.M., Hoffman, S.L., Gardner, M.J., Carucci, D.J., 2002. Genome sequence and comparative analysis of the model rodent malaria parasite *Plasmodium yoelii yoelii*. *Nature* 419, 512-519.
- Carter, R., Walliker, D., 1975. New observations on the malaria parasites of rodents of the Central African Republic - *Plasmodium vinckei petteri* subsp. nov. and *Plasmodium chabaudi* Landau, 1965. *Ann Trop Med Parasitol* 69, 187-196.
- Carter, R., 1978. Studies on enzyme variation in the murine malaria parasites *Plasmodium berghei*, *P. yoelii*, *P. vinckei* and *P. chabaudi* by starch gel electrophoresis. *Parasitology* 76, 241-267.
- Carter, R., Hunt, P., Cheesman, S., 2007. Linkage Group Selection--a fast approach to the genetic analysis of malaria parasites. *Int J Parasitol* 37, 285-293.
- Carver, T., Berriman, M., Tivey, A., Patel, C., Bohme, U., Barrell, B.G., Parkhill, J., Rajandream, M.A., 2008. Artemis and ACT: viewing, annotating and comparing sequences stored in a relational database. *Bioinformatics* 24, 2672-2676.
- Chavchich, M., Gerena, L., Peters, J., Chen, N., Cheng, Q., Kyle, D.E., 2010. Role of *pfmdr1* amplification and expression in induction of resistance to artemisinin derivatives in *Plasmodium falciparum*. *Antimicrob Agents Chemother* 54, 2455-2464.
- Chawira, A.N., Warhurst, D.C., Peters, W., 1986. Qinghaosu resistance in rodent malaria. *Trans R Soc Trop Med Hyg* 80, 477-480.

- Cheesman, S., Creasey, A., Degnan, K., Kooij, T., Afonso, A., Cravo, P., Carter, R., Hunt, P., 2007. Validation of Pyrosequencing for accurate and high throughput estimation of allele frequencies in malaria parasites. *Mol Biochem Parasitol* 152, 213-219.
- Cheesman, S., O'Mahony, E., Pattaradilokrat, S., Degnan, K., Knott, S., Carter, R., 2010a. A single parasite gene determines strain-specific protective immunity against malaria: The role of the merozoite surface protein I. *Int J Parasitol*.
- Cheesman, S., O'Mahony, E., Pattaradilokrat, S., Degnan, K., Knott, S., Carter, R., 2010b. A single parasite gene determines strain-specific protective immunity against malaria: the role of the merozoite surface protein I. *Int J Parasitol* 40, 951-961.
- Chen, N., Baker, J., Ezard, N., Burns, M., Edstein, M.D., Cheng, Q., 2002. Short report: Molecular evaluation of the efficacy of chloroquine treatment of uncomplicated *Plasmodium falciparum* malaria in East Timor. *Am J Trop Med Hyg* 67, 64-66.
- Chen, N., Kyle, D.E., Pasay, C., Fowler, E.V., Baker, J., Peters, J.M., Cheng, Q., 2003. pfcrt Allelic types with two novel amino acid mutations in chloroquine-resistant *Plasmodium falciparum* isolates from the Philippines. *Antimicrob Agents Chemother* 47, 3500-3505.
- Cheng, Q., Saul, A., 1994. The dihydrofolate reductase domain of rodent malarias: point mutations and pyrimethamine resistance. *Mol Biochem Parasitol* 65, 361-363.
- Chevalley, S., Coste, A., Lopez, A., Pipy, B., Valentin, A., 2010. Flow cytometry for the evaluation of anti-plasmodial activity of drugs on *Plasmodium falciparum* gametocytes. *Malar J* 9, 49.
- Clyde, D.F., Shute, G.T., 1959. Survival of pyrimethamine-resistant *Plasmodium falciparum*. *Trans R Soc Trop Med Hyg* 53, 170-172.
- Coatney, G.R., 1963. Pitfalls in a discovery: the chronicle of chloroquine. *Am J Trop Med Hyg* 12, 121-128.
- Cogswell, F.B., 1992. The hypnozoite and relapse in primate malaria. *Clin Microbiol Rev* 5, 26-35.
- Conway, D.J., 2007. Molecular epidemiology of malaria. *Clin Microbiol Rev* 20, 188-204.
- Cowman, A.F., Lew, A.M., 1990. Chromosomal rearrangements and point mutations in the DHFR-TS gene of *Plasmodium chabaudi* under antifolate selection. *Mol Biochem Parasitol* 42, 21-29.
- Cowman, A.F., Karcz, S., Galatis, D., Culvenor, J.G., 1991. A P-glycoprotein homologue of *Plasmodium falciparum* is localized on the digestive vacuole. *J Cell Biol* 113, 1033-1042.
- Cowman, A.F., Galatis, D., Thompson, J.K., 1994. Selection for mefloquine resistance in *Plasmodium falciparum* is linked to amplification of the pfmdr1 gene and cross-resistance to halofantrine and quinine. *Proc Natl Acad Sci U S A* 91, 1143-1147.
- Cox-Singh, J., Davis, T.M., Lee, K.S., Shamsul, S.S., Matusop, A., Ratnam, S., Rahman, H.A., Conway, D.J., Singh, B., 2008. *Plasmodium knowlesi* malaria in humans is widely distributed and potentially life threatening. *Clin Infect Dis* 46, 165-171.

- Cravo, P.V., Carlton, J.M., Hunt, P., Bisoni, L., Padua, R.A., Walliker, D., 2003. Genetics of mefloquine resistance in the rodent malaria parasite *Plasmodium chabaudi*. *Antimicrob Agents Chemother* 47, 709-718.
- Culleton, R., Martinelli, A., Hunt, P., Carter, R., 2005. Linkage group selection: rapid gene discovery in malaria parasites. *Genome Res* 15, 92-97.
- Cunningham, D., Lawton, J., Jarra, W., Preiser, P., Langhorne, J., 2010. The *pir* multigene family of *Plasmodium*: antigenic variation and beyond. *Mol Biochem Parasitol* 170, 65-73.
- d'Azzo, A., Bongiovanni, A., Nastasi, T., 2005. E3 ubiquitin ligases as regulators of membrane protein trafficking and degradation. *Traffic* 6, 429-441.
- Dahl, E.L., Rosenthal, P.J., 2007. Multiple antibiotics exert delayed effects against the *Plasmodium falciparum* apicoplast. *Antimicrob Agents Chemother* 51, 3485-3490.
- Darwin, C., 1859. *On the Origin of Species*.
- de Roode, J.C., Helinski, M.E., Anwar, M.A., Read, A.F., 2005. Dynamics of multiple infection and within-host competition in genetically diverse malaria infections. *Am Nat* 166, 531-542.
- de Souza, J.B., Riley, E.M., 2002. Cerebral malaria: the contribution of studies in animal models to our understanding of immunopathogenesis. *Microbes Infect* 4, 291-300.
- Denis, M.B., Tsuyuoka, R., Poravuth, Y., Narann, T.S., Seila, S., Lim, C., Incardona, S., Lim, P., Sem, R., Socheat, D., Christophel, E.M., Ringwald, P., 2006. Surveillance of the efficacy of artesunate and mefloquine combination for the treatment of uncomplicated falciparum malaria in Cambodia. *Trop Med Int Health* 11, 1360-1366.
- Djimde, A., Doumbo, O.K., Cortese, J.F., Kayentao, K., Doumbo, S., Diourte, Y., Dicko, A., Su, X.Z., Nomura, T., Fidock, D.A., Wellems, T.E., Plowe, C.V., Coulibaly, D., 2001. A molecular marker for chloroquine-resistant falciparum malaria. *N Engl J Med* 344, 257-263.
- Dondorp, A.M., Nosten, F., Yi, P., Das, D., Phyto, A.P., Tarning, J., Lwin, K.M., Arie, F., Hanpithakpong, W., Lee, S.J., Ringwald, P., Silamut, K., Imwong, M., Chotivanich, K., Lim, P., Herdman, T., An, S.S., Yeung, S., Singhasivanon, P., Day, N.P., Lindegardh, N., Socheat, D., White, N.J., 2009. Artemisinin resistance in *Plasmodium falciparum* malaria. *N Engl J Med* 361, 455-467.
- Dondorp, A.M., Yeung, S., White, L., Nguon, C., Day, N.P., Socheat, D., von Seidlein, L., 2010. Artemisinin resistance: current status and scenarios for containment. *Nat Rev Microbiol* 8, 272-280.
- Duraisingh, M.T., von Seidlein, L.V., Jepson, A., Jones, P., Sambou, I., Pinder, M., Warhurst, D.C., 2000. Linkage disequilibrium between two chromosomally distinct loci associated with increased resistance to chloroquine in *Plasmodium falciparum*. *Parasitology* 121 (Pt 1), 1-7.
- Duraisingh, M.T., Cowman, A.F., 2005. Contribution of the *pfmdr1* gene to antimalarial drug-resistance. *Acta Trop* 94, 181-190.
- Durand, R., Jafari, S., Vauzelle, J., Delabre, J.F., Jesic, Z., Le Bras, J., 2001. Analysis of *pfert* point mutations and chloroquine susceptibility in isolates of *Plasmodium falciparum*. *Mol Biochem Parasitol* 114, 95-102.

- Durrand, V., Berry, A., Sem, R., Glaziou, P., Beaudou, J., Fandeur, T., 2004. Variations in the sequence and expression of the *Plasmodium falciparum* chloroquine resistance transporter (PfCRT) and their relationship to chloroquine resistance in vitro. *Mol Biochem Parasitol* 136, 273-285.
- Eckstein-Ludwig, U., Webb, R.J., Van Goethem, I.D., East, J.M., Lee, A.G., Kimura, M., O'Neill, P.M., Bray, P.G., Ward, S.A., Krishna, S., 2003. Artemisinins target the SERCA of *Plasmodium falciparum*. *Nature* 424, 957-961.
- Ehrenreich, I.M., Torabi, N., Jia, Y., Kent, J., Martis, S., Shapiro, J.A., Gresham, D., Caudy, A.A., Kruglyak, L., 2010. Dissection of genetically complex traits with extremely large pools of yeast segregants. *Nature* 464, 1039-1042.
- Enayati, A., Hemingway, J., 2010. Malaria management: past, present, and future. *Annu Rev Entomol* 55, 569-591.
- Famin, O., Ginsburg, H., 2002. Differential effects of 4-aminoquinoline-containing antimalarial drugs on hemoglobin digestion in *Plasmodium falciparum*-infected erythrocytes. *Biochem Pharmacol* 63, 393-398.
- Fidock, D.A., Nomura, T., Talley, A.K., Cooper, R.A., Dzekunov, S.M., Ferdig, M.T., Ursos, L.M., Sidhu, A.B., Naude, B., Deitsch, K.W., Su, X.Z., Wootton, J.C., Roepe, P.D., Wellems, T.E., 2000. Mutations in the *P. falciparum* digestive vacuole transmembrane protein PfCRT and evidence for their role in chloroquine resistance. *Mol Cell* 6, 861-871.
- Foote, S.J., Thompson, J.K., Cowman, A.F., Kemp, D.J., 1989. Amplification of the multidrug resistance gene in some chloroquine-resistant isolates of *P. falciparum*. *Cell* 57, 921-930.
- Foote, S.J., Kyle, D.E., Martin, R.K., Oduola, A.M., Forsyth, K., Kemp, D.J., Cowman, A.F., 1990. Several alleles of the multidrug-resistance gene are closely linked to chloroquine resistance in *Plasmodium falciparum*. *Nature* 345, 255-258.
- Galinski, M.R., Barnwell, J.W., 2008. *Plasmodium vivax*: who cares? *Malar J* 7 Suppl 1, S9.
- Gervais, G.W., Trujillo, K., Robinson, B.L., Peters, W., Serrano, A.E., 1999. *Plasmodium berghei*: identification of an mdr-like gene associated with drug resistance. *Exp Parasitol* 91, 86-92.
- Ginsburg, H., 2005. Should chloroquine be laid to rest? *Acta Trop* 96, 16-23.
- Girard, M.P., Reed, Z.H., Friede, M., Kieny, M.P., 2007. A review of human vaccine research and development: malaria. *Vaccine* 25, 1567-1580.
- Gligorijevic, B., Bennett, T., McAllister, R., Urbach, J.S., Roepe, P.D., 2006. Spinning disk confocal microscopy of live, intraerythrocytic malarial parasites. 2. Altered vacuolar volume regulation in drug resistant malaria. *Biochemistry* 45, 12411-12423.
- Gligorijevic, B., Purdy, K., Elliott, D.A., Cooper, R.A., Roepe, P.D., 2008. Stage independent chloroquine resistance and chloroquine toxicity revealed via spinning disk confocal microscopy. *Mol Biochem Parasitol* 159, 7-23.
- Goldberg, D.E., Slater, A.F., 1992. The pathway of hemoglobin degradation in malaria parasites. *Parasitol Today* 8, 280-283.
- Grech, K., Martinelli, A., Pathirana, S., Walliker, D., Hunt, P., Carter, R., 2002. Numerous, robust genetic markers for *Plasmodium chabaudi* by the method of amplified fragment length polymorphism. *Mol Biochem Parasitol* 123, 95-104.

- Gwer, S., Newton, C.R., Berkley, J.A., 2007. Over-diagnosis and co-morbidity of severe malaria in African children: a guide for clinicians. *Am J Trop Med Hyg* 77, 6-13.
- Hall, N., Karras, M., Raine, J.D., Carlton, J.M., Kooij, T.W., Berriman, M., Florens, L., Janssen, C.S., Pain, A., Christophides, G.K., James, K., Rutherford, K., Harris, B., Harris, D., Churcher, C., Quail, M.A., Ormond, D., Doggett, J., Trueman, H.E., Mendoza, J., Bidwell, S.L., Rajandream, M.A., Carucci, D.J., Yates, J.R., 3rd, Kafatos, F.C., Janse, C.J., Barrell, B., Turner, C.M., Waters, A.P., Sinden, R.E., 2005. A comprehensive survey of the *Plasmodium* life cycle by genomic, transcriptomic, and proteomic analyses. *Science* 307, 82-86.
- Harinasuta, T., Suntharasamai, P., Viravan, C., 1965. Chloroquine-resistant falciparum malaria in Thailand. *Lancet* 2, 657-660.
- Hastings, I.M., 2004. The origins of antimalarial drug resistance. *Trends Parasitol* 20, 512-518.
- Hastings, I.M., Donnelly, M.J., 2005. The impact of antimalarial drug resistance mutations on parasite fitness, and its implications for the evolution of resistance. *Drug Resist Updat* 8, 43-50.
- Hayton, K., Ranford-Cartwright, L.C., Walliker, D., 2002. Sulfadoxine-pyrimethamine resistance in the rodent malaria parasite *Plasmodium chabaudi*. *Antimicrob Agents Chemother* 46, 2482-2489.
- Hayton, K., Gaur, D., Liu, A., Takahashi, J., Henschen, B., Singh, S., Lambert, L., Furuya, T., Bouttenot, R., Doll, M., Nawaz, F., Mu, J., Jiang, L., Miller, L.H., Wellems, T.E., 2008. Erythrocyte binding protein PfRH5 polymorphisms determine species-specific pathways of *Plasmodium falciparum* invasion. *Cell Host Microbe* 4, 40-51.
- Hayton, K., Su, X.Z., 2008. Drug resistance and genetic mapping in *Plasmodium falciparum*. *Curr Genet* 54, 223-239.
- Hayward, R., Saliba, K.J., Kirk, K., 2005. *pfmdr1* mutations associated with chloroquine resistance incur a fitness cost in *Plasmodium falciparum*. *Mol Microbiol* 55, 1285-1295.
- Hong, Y.L., Yang, Y.Z., Meshnick, S.R., 1994. The interaction of artemisinin with malarial hemozoin. *Mol Biochem Parasitol* 63, 121-128.
- Hu, M., Li, P., Li, M., Li, W., Yao, T., Wu, J.W., Gu, W., Cohen, R.E., Shi, Y., 2002. Crystal structure of a UBP-family deubiquitinating enzyme in isolation and in complex with ubiquitin aldehyde. *Cell* 111, 1041-1054.
- Hughes, A.L., Verra, F., 2010. Malaria Parasite Sequences from Chimpanzee Support the Co-Speciation Hypothesis for the Origin of Virulent Human Malaria (*Plasmodium falciparum*). *Mol Phylogenet Evol*.
- Hunt, P., Cravo, P.V., Donleavy, P., Carlton, J.M., Walliker, D., 2004a. Chloroquine resistance in *Plasmodium chabaudi*: are chloroquine-resistance transporter (*crt*) and multi-drug resistance (*mdr1*) orthologues involved? *Mol Biochem Parasitol* 133, 27-35.
- Hunt, P., Martinelli, A., Fawcett, R., Carlton, J., Carter, R., Walliker, D., 2004b. Gene synteny and chloroquine resistance in *Plasmodium chabaudi*. *Mol Biochem Parasitol* 136, 157-164.

- Hunt, P., Fawcett, R., Carter, R., Walliker, D., 2005. Estimating SNP proportions in populations of malaria parasites by sequencing: validation and applications. *Mol Biochem Parasitol* 143, 173-182.
- Hunt, P., Afonso, A., Creasey, A., Culleton, R., Sidhu, A.B., Logan, J., Valderramos, S.G., McNae, I., Cheesman, S., do Rosario, V., Carter, R., Fidock, D.A., Cravo, P., 2007. Gene encoding a deubiquitinating enzyme is mutated in artesunate- and chloroquine-resistant rodent malaria parasites. *Mol Microbiol* 65, 27-40.
- Hyde, J.E., 2005. Exploring the folate pathway in *Plasmodium falciparum*. *Acta Trop* 94, 191-206.
- Hyde, J.E., Dittrich, S., Wang, P., Sims, P.F., de Crecy-Lagard, V., Hanson, A.D., 2008. *Plasmodium falciparum*: a paradigm for alternative folate biosynthesis in diverse microorganisms? *Trends Parasitol* 24, 502-508.
- Imwong, M., Dondorp, A.M., Nosten, F., Yi, P., Mungthin, M., Hanchana, S., Das, D., Phyo, A.P., Lwin, K.M., Pukrittayakamee, S., Lee, S.J., Saisung, S., Koecharoen, K., Nguon, C., Day, N.P., Socheat, D., White, N.J., 2010. Exploring the contribution of candidate genes to artemisinin resistance in *Plasmodium falciparum*. *Antimicrob Agents Chemother* 54, 2886-2892.
- Jambou, R., Legrand, E., Niang, M., Khim, N., Lim, P., Volney, B., Ekala, M.T., Bouchier, C., Esterre, P., Fandeur, T., Mercereau-Puijalon, O., 2005. Resistance of *Plasmodium falciparum* field isolates to in-vitro artemether and point mutations of the SERCA-type PfATPase6. *Lancet* 366, 1960-1963.
- Janssen, C.S., Phillips, R.S., Turner, C.M., Barrett, M.P., 2004. *Plasmodium* interspersed repeats: the major multigene superfamily of malaria parasites. *Nucleic Acids Res* 32, 5712-5720.
- Jearnpipatkul, A., Govitrapong, P., Yuthavong, Y., Wilairat, P., Panijpan, B., 1980. Binding of antimalarial drugs to hemozoin from *Plasmodium berghei*. *Experientia* 36, 1063-1064.
- Jeffares, D.C., Pain, A., Berry, A., Cox, A.V., Stalker, J., Ingle, C.E., Thomas, A., Quail, M.A., Siebenthall, K., Uhlemann, A.C., Kyes, S., Krishna, S., Newbold, C., Dermitzakis, E.T., Berriman, M., 2007. Genome variation and evolution of the malaria parasite *Plasmodium falciparum*. *Nat Genet* 39, 120-125.
- Jeffery, G.M., 1976. Malaria control in the twentieth century. *Am J Trop Med Hyg* 25, 361-371.
- Jiang, H., Patel, J.J., Yi, M., Mu, J., Ding, J., Stephens, R., Cooper, R.A., Ferdig, M.T., Su, X.Z., 2008. Genome-wide compensatory changes accompany drug-selected mutations in the *Plasmodium falciparum* crt gene. *PLoS One* 3, e2484.
- Johnson, D.J., Fidock, D.A., Mungthin, M., Lakshmanan, V., Sidhu, A.B., Bray, P.G., Ward, S.A., 2004. Evidence for a central role for PfCRT in conferring *Plasmodium falciparum* resistance to diverse antimalarial agents. *Mol Cell* 15, 867-877.
- Juliano, R.L., Ling, V., 1976. A surface glycoprotein modulating drug permeability in Chinese hamster ovary cell mutants. *Biochim Biophys Acta* 455, 152-162.
- Jung, M., Kim, H., Nam, K.Y., No, K.T., 2005. Three-dimensional structure of *Plasmodium falciparum* Ca²⁺-ATPase(PfATP6) and docking of artemisinin derivatives to PfATP6. *Bioorg Med Chem Lett* 15, 2994-2997.

- Kallander, K., Nsungwa-Sabiiti, J., Peterson, S., 2004. Symptom overlap for malaria and pneumonia--policy implications for home management strategies. *Acta Trop* 90, 211-214.
- Kiszewski, A., Mellinger, A., Spielman, A., Malaney, P., Sachs, S.E., Sachs, J., 2004. A global index representing the stability of malaria transmission. *Am J Trop Med Hyg* 70, 486-498.
- Kooij, T.W., Carlton, J.M., Bidwell, S.L., Hall, N., Ramesar, J., Janse, C.J., Waters, A.P., 2005. A *Plasmodium* whole-genome synteny map: indels and synteny breakpoints as foci for species-specific genes. *PLoS Pathog* 1, e44.
- Krogstad, D.J., Gluzman, I.Y., Kyle, D.E., Oduola, A.M., Martin, S.K., Milhous, W.K., Schlesinger, P.H., 1987. Efflux of chloroquine from *Plasmodium falciparum*: mechanism of chloroquine resistance. *Science* 238, 1283-1285.
- Kublin, J.G., Cortese, J.F., Njunju, E.M., Mukadam, R.A., Wirima, J.J., Kazembe, P.N., Djimde, A.A., Kouriba, B., Taylor, T.E., Plowe, C.V., 2003. Reemergence of chloroquine-sensitive *Plasmodium falciparum* malaria after cessation of chloroquine use in Malawi. *J Infect Dis* 187, 1870-1875.
- LaCount, D.J., Vignali, M., Chettier, R., Phansalkar, A., Bell, R., Hesselberth, J.R., Schoenfeld, L.W., Ota, I., Sahasrabudhe, S., Kurschner, C., Fields, S., Hughes, R.E., 2005. A protein interaction network of the malaria parasite *Plasmodium falciparum*. *Nature* 438, 103-107.
- Lakshmanan, V., Bray, P.G., Verdier-Pinard, D., Johnson, D.J., Horrocks, P., Muhle, R.A., Alakpa, G.E., Hughes, R.H., Ward, S.A., Krogstad, D.J., Sidhu, A.B., Fidock, D.A., 2005. A critical role for PfCRT K76T in *Plasmodium falciparum* verapamil-reversible chloroquine resistance. *Embo J* 24, 2294-2305.
- Lamikanra, A.A., Brown, D., Potocnik, A., Casals-Pascual, C., Langhorne, J., Roberts, D.J., 2007. Malarial anemia: of mice and men. *Blood* 110, 18-28.
- Laufer, M.K., Thesing, P.C., Eddington, N.D., Masonga, R., Dzinjalama, F.K., Takala, S.L., Taylor, T.E., Plowe, C.V., 2006. Return of chloroquine antimalarial efficacy in Malawi. *N Engl J Med* 355, 1959-1966.
- Le Roch, K.G., Zhou, Y., Blair, P.L., Grainger, M., Moch, J.K., Haynes, J.D., De La Vega, P., Holder, A.A., Batalov, S., Carucci, D.J., Winzeler, E.A., 2003. Discovery of gene function by expression profiling of the malaria parasite life cycle. *Science* 301, 1503-1508.
- Lengeler, C., Snow, R.W., 1996. From efficacy to effectiveness: insecticide-treated bednets in Africa. *Bull World Health Organ* 74, 325-332.
- Li, C., Seixas, E., Langhorne, J., 2001. Rodent malarias: the mouse as a model for understanding immune responses and pathology induced by the erythrocytic stages of the parasite. *Med Microbiol Immunol* 189, 115-126.
- Li, H., Ruan, J., Durbin, R., 2008. Mapping short DNA sequencing reads and calling variants using mapping quality scores. *Genome Res* 18, 1851-1858.
- Li, W., Mo, W., Shen, D., Sun, L., Wang, J., Lu, S., Gitschier, J.M., Zhou, B., 2005. Yeast model uncovers dual roles of mitochondria in action of artemisinin. *PLoS Genet* 1, e36.
- Lim, P., Chy, S., Arie, F., Incardona, S., Chim, P., Sem, R., Denis, M.B., Hewitt, S., Hoyer, S., Socheat, D., Merecreau-Puijalon, O., Fandeur, T., 2003. pfcrt polymorphism and chloroquine resistance in *Plasmodium falciparum* strains isolated in Cambodia. *Antimicrob Agents Chemother* 47, 87-94.

- Liu, S., Chen, H.D., Makarevitch, I., Shirmer, R., Emrich, S.J., Dietrich, C.R., Barbazuk, W.B., Springer, N.M., Schnable, P.S., 2010. High-throughput genetic mapping of mutants via quantitative single nucleotide polymorphism typing. *Genetics* 184, 19-26.
- Lozovsky, E.R., Chookajorn, T., Brown, K.M., Imwong, M., Shaw, P.J., Kamchonwongpaisan, S., Neafsey, D.E., Weinreich, D.M., Hartl, D.L., 2009. Stepwise acquisition of pyrimethamine resistance in the malaria parasite. *Proc Natl Acad Sci U S A* 106, 12025-12030.
- Lubell, Y., Reyburn, H., Mbakilwa, H., Mwangi, R., Chonya, K., Whitty, C.J., Mills, A., 2007. The cost-effectiveness of parasitologic diagnosis for malaria-suspected patients in an era of combination therapy. *Am J Trop Med Hyg* 77, 128-132.
- Marshall, J.M., Taylor, C.E., 2009. Malaria control with transgenic mosquitoes. *PLoS Med* 6, e20.
- Martin, R.E., Kirk, K., 2004. The malaria parasite's chloroquine resistance transporter is a member of the drug/metabolite transporter superfamily. *Mol Biol Evol* 21, 1938-1949.
- Martin, R.E., Henry, R.I., Abbey, J.L., Clements, J.D., Kirk, K., 2005. The 'permeome' of the malaria parasite: an overview of the membrane transport proteins of *Plasmodium falciparum*. *Genome Biol* 6, R26.
- Martin, R.E., Marchetti, R.V., Cowan, A.I., Howitt, S.M., Broer, S., Kirk, K., 2009. Chloroquine transport via the malaria parasite's chloroquine resistance transporter. *Science* 325, 1680-1682.
- Martin, S.K., Oduola, A.M., Milhous, W.K., 1987. Reversal of chloroquine resistance in *Plasmodium falciparum* by verapamil. *Science* 235, 899-901.
- Martinelli, A., Hunt, P., Cheesman, S.J., Carter, R., 2004. Amplified fragment length polymorphism measures proportions of malaria parasites carrying specific alleles in complex genetic mixtures. *Mol Biochem Parasitol* 136, 117-122.
- Martinelli, A., Cheesman, S., Hunt, P., Culleton, R., Raza, A., Mackinnon, M., Carter, R., 2005. A genetic approach to the de novo identification of targets of strain-specific immunity in malaria parasites. *Proc Natl Acad Sci U S A* 102, 814-819.
- Mather, M.W., Darrouzet, E., Valkova-Valchanova, M., Cooley, J.W., McIntosh, M.T., Daldal, F., Vaidya, A.B., 2005. Uncovering the molecular mode of action of the antimalarial drug atovaquone using a bacterial system. *J Biol Chem* 280, 27458-27465.
- Mendis, K., Sina, B.J., Marchesini, P., Carter, R., 2001. The neglected burden of *Plasmodium vivax* malaria. *Am J Trop Med Hyg* 64, 97-106.
- Meshnick, S.R., 1994. The mode of action of antimalarial endoperoxides. *Trans R Soc Trop Med Hyg* 88 Suppl 1, S31-32.
- Meshnick, S.R., Jefford, C.W., Posner, G.H., Avery, M.A., Peters, W., 1996a. Second-generation antimalarial endoperoxides. *Parasitol Today* 12, 79-82.
- Meshnick, S.R., Taylor, T.E., Kamchonwongpaisan, S., 1996b. Artemisinin and the antimalarial endoperoxides: from herbal remedy to targeted chemotherapy. *Microbiol Rev* 60, 301-315.
- Miki, A., Tanabe, K., Nakayama, T., Kiryon, C., Ohsawa, K., 1992. *Plasmodium chabaudi*: association of reversal of chloroquine resistance with increased accumulation of chloroquine in resistant parasites. *Exp Parasitol* 74, 134-142.

- Mita, T., Kaneko, A., Lum, J.K., Bwijo, B., Takechi, M., Zungu, I.L., Tsukahara, T., Tanabe, K., Kobayakawa, T., Bjorkman, A., 2003. Recovery of chloroquine sensitivity and low prevalence of the *Plasmodium falciparum* chloroquine resistance transporter gene mutation K76T following the discontinuance of chloroquine use in Malawi. *Am J Trop Med Hyg* 68, 413-415.
- Mita, T., Kaneko, A., Hombhanje, F., Hwaihwanje, I., Takahashi, N., Osawa, H., Tsukahara, T., Masta, A., Lum, J.K., Kobayakawa, T., Ishizaki, T., Bjorkman, A., 2006. Role of *pfmdr1* mutations on chloroquine resistance in *Plasmodium falciparum* isolates with *pfcr* K76T from Papua New Guinea. *Acta Trop* 98, 137-144.
- Mita, T., Tanabe, K., Kita, K., 2009. Spread and evolution of *Plasmodium falciparum* drug resistance. *Parasitol Int* 58, 201-209.
- Moore, D.V., Lanier, J.E., 1961. Observations on two *Plasmodium falciparum* infections with an abnormal response to chloroquine. *Am J Trop Med Hyg* 10, 5-9.
- Mu, J., Duan, J., Makova, K.D., Joy, D.A., Huynh, C.Q., Branch, O.H., Li, W.H., Su, X.Z., 2002. Chromosome-wide SNPs reveal an ancient origin for *Plasmodium falciparum*. *Nature* 418, 323-326.
- Mu, J., Awadalla, P., Duan, J., McGee, K.M., Keebler, J., Seydel, K., McVean, G.A., Su, X.Z., 2007. Genome-wide variation and identification of vaccine targets in the *Plasmodium falciparum* genome. *Nat Genet* 39, 126-130.
- Mu, J., Myers, R.A., Jiang, H., Liu, S., Ricklefs, S., Waisberg, M., Chotivanich, K., Wilairatana, P., Krudsood, S., White, N.J., Udomsangpetch, R., Cui, L., Ho, M., Ou, F., Li, H., Song, J., Li, G., Wang, X., Seila, S., Sokunthea, S., Socheat, D., Sturdevant, D.E., Porcella, S.F., Fairhurst, R.M., Wellems, T.E., Awadalla, P., Su, X.Z., 2010. *Plasmodium falciparum* genome-wide scans for positive selection, recombination hot spots and resistance to antimalarial drugs. *Nat Genet* 42, 268-271.
- Nagesha, H.S., Din, S., Casey, G.J., Susanti, A.I., Fryauff, D.J., Reeder, J.C., Cowman, A.F., 2001. Mutations in the *pfmdr1*, *dhfr* and *dhps* genes of *Plasmodium falciparum* are associated with in-vivo drug resistance in West Papua, Indonesia. *Trans R Soc Trop Med Hyg* 95, 43-49.
- Nagesha, H.S., Casey, G.J., Rieckmann, K.H., Fryauff, D.J., Laksana, B.S., Reeder, J.C., Maguire, J.D., Baird, J.K., 2003. New haplotypes of the *Plasmodium falciparum* chloroquine resistance transporter (*pfcr*) gene among chloroquine-resistant parasite isolates. *Am J Trop Med Hyg* 68, 398-402.
- Nair, S., Williams, J.T., Brockman, A., Paiphun, L., Mayxay, M., Newton, P.N., Guthmann, J.P., Smithuis, F.M., Hien, T.T., White, N.J., Nosten, F., Anderson, T.J., 2003. A selective sweep driven by pyrimethamine treatment in southeast asian malaria parasites. *Mol Biol Evol* 20, 1526-1536.
- Nair, S., Nash, D., Sudimack, D., Jaidee, A., Barends, M., Uhlemann, A.C., Krishna, S., Nosten, F., Anderson, T.J., 2007. Recurrent gene amplification and soft selective sweeps during evolution of multidrug resistance in malaria parasites. *Mol Biol Evol* 24, 562-573.
- Naude, B., Brzostowski, J.A., Kimmel, A.R., Wellems, T.E., 2005. *Dictyostelium discoideum* expresses a malaria chloroquine resistance mechanism upon transfection with mutant, but not wild-type, *Plasmodium falciparum* transporter PfCRT. *J Biol Chem* 280, 25596-25603.

- Ning, Z., Cox, A.J., Mullikin, J.C., 2001. SSAHA: a fast search method for large DNA databases. *Genome Res* 11, 1725-1729.
- Noedl, H., 2009. ABC - antibiotics-based combinations for the treatment of severe malaria? *Trends Parasitol* 25, 540-544.
- Noedl, H., Socheat, D., Satimai, W., 2009. Artemisinin-resistant malaria in Asia. *N Engl J Med* 361, 540-541.
- Nomura, T., Carlton, J.M., Baird, J.K., del Portillo, H.A., Fryauff, D.J., Rathore, D., Fidock, D.A., Su, X., Collins, W.E., McCutchan, T.F., Wootton, J.C., Wellems, T.E., 2001. Evidence for different mechanisms of chloroquine resistance in 2 *Plasmodium* species that cause human malaria. *J Infect Dis* 183, 1653-1661.
- Nzila, A., 2006. The past, present and future of antifolates in the treatment of *Plasmodium falciparum* infection. *J Antimicrob Chemother* 57, 1043-1054.
- Nzila, A., Mwai, L., 2009. In vitro selection of *Plasmodium falciparum* drug-resistant parasite lines. *J Antimicrob Chemother* 65, 390-398.
- Ohsawa, K., Tanabe, K., Kimata, I., Miki, A., 1991. Ultrastructural changes associated with reversal of chloroquine resistance by verapamil in *Plasmodium chabaudi*. *Parasitology* 103 Pt 2, 185-189.
- Olliaro, P., Wells, T.N., 2009. The global portfolio of new antimalarial medicines under development. *Clin Pharmacol Ther* 85, 584-595.
- Ord, R., Alexander, N., Dunyo, S., Hallett, R., Jawara, M., Targett, G., Drakeley, C.J., Sutherland, C.J., 2007. Seasonal carriage of *pfcr*t and *pfmdr*1 alleles in Gambian *Plasmodium falciparum* imply reduced fitness of chloroquine-resistant parasites. *J Infect Dis* 196, 1613-1619.
- Orjuela-Sanchez, P., de Santana Filho, F.S., Machado-Lima, A., Chehuan, Y.F., Costa, M.R., Alecrim, M.G., del Portillo, H.A., 2009. Analysis of single-nucleotide polymorphisms in the *crt-o* and *mdr1* genes of *Plasmodium vivax* among chloroquine-resistant isolates from the Brazilian Amazon region. *Antimicrob Agents Chemother* 53, 3561-3564.
- Padua, R.A., 1981. *Plasmodium chabaudi*: genetics of resistance to chloroquine. *Exp Parasitol* 52, 419-426.
- Paguio, M.F., Cabrera, M., Roepe, P.D., 2009. Chloroquine transport in *Plasmodium falciparum*. 2. Analysis of PfCRT-mediated drug transport using proteoliposomes and a fluorescent chloroquine probe. *Biochemistry* 48, 9482-9491.
- Patel, J.J., Thacker, D., Tan, J.C., Pleeter, P., Checkley, L., Gonzales, J.M., Deng, B., Roepe, P.D., Cooper, R.A., Ferdig, M.T., 2010. Chloroquine susceptibility and reversibility in a *Plasmodium falciparum* genetic cross. *Mol Microbiol* 78, 770-787.
- Pattaradilokrat, S., Culleton, R.L., Cheesman, S.J., Carter, R., 2009. Gene encoding erythrocyte binding ligand linked to blood stage multiplication rate phenotype in *Plasmodium yoelii yoelii*. *Proc Natl Acad Sci U S A* 106, 7161-7166.
- Peel, S.A., Bright, P., Yount, B., Handy, J., Baric, R.S., 1994. A strong association between mefloquine and halofantrine resistance and amplification, overexpression, and mutation in the P-glycoprotein gene homolog (*pfmdr*) of *Plasmodium falciparum* in vitro. *Am J Trop Med Hyg* 51, 648-658.
- Peters, J.M., Chen, N., Gatton, M., Korsinczyk, M., Fowler, E.V., Manzetti, S., Saul, A., Cheng, Q., 2002. Mutations in cytochrome b resulting in atovaquone

- resistance are associated with loss of fitness in *Plasmodium falciparum*. *Antimicrob Agents Chemother* 46, 2435-2441.
- Peterson, D.S., Walliker, D., Wellems, T.E., 1988. Evidence that a point mutation in dihydrofolate reductase-thymidylate synthase confers resistance to pyrimethamine in *falciparum* malaria. *Proc Natl Acad Sci U S A* 85, 9114-9118.
- Phan, G.T., de Vries, P.J., Tran, B.Q., Le, H.Q., Nguyen, N.V., Nguyen, T.V., Heisterkamp, S.H., Kager, P.A., 2002. Artemisinin or chloroquine for blood stage *Plasmodium vivax* malaria in Vietnam. *Trop Med Int Health* 7, 858-864.
- Pierce, S.K., Miller, L.H., 2009. World Malaria Day 2009: what malaria knows about the immune system that immunologists still do not. *J Immunol* 182, 5171-5177.
- Poudel, S.S., Newman, R.A., Vaughan, J.A., 2008. Rodent *Plasmodium*: population dynamics of early sporogony within *Anopheles stephensi* mosquitoes. *J Parasitol* 94, 999-1008.
- Powers, K.G., Jacobs, R.L., Good, W.C., Koontz, L.C., 1969. *Plasmodium vinckei*: production of chloroquine-resistant strain. *Exp Parasitol* 26, 193-202.
- Price, R.N., Uhlemann, A.C., Brockman, A., McGready, R., Ashley, E., Phaipun, L., Patel, R., Laing, K., Looareesuwan, S., White, N.J., Nosten, F., Krishna, S., 2004. Mefloquine resistance in *Plasmodium falciparum* and increased *pfmdr1* gene copy number. *Lancet* 364, 438-447.
- Price, R.N., Tjitra, E., Guerra, C.A., Yeung, S., White, N.J., Anstey, N.M., 2007. *Vivax* malaria: neglected and not benign. *Am J Trop Med Hyg* 77, 79-87.
- Prudencio, M., Rodriguez, A., Mota, M.M., 2006. The silent path to thousands of merozoites: the *Plasmodium* liver stage. *Nat Rev Microbiol* 4, 849-856.
- Pukrittayakamee, S., Chotivanich, K., Chantira, A., Clemens, R., Looareesuwan, S., White, N.J., 2004. Activities of artesunate and primaquine against asexual- and sexual-stage parasites in *falciparum* malaria. *Antimicrob Agents Chemother* 48, 1329-1334.
- Puri, S.K., Chandra, R., 2006. *Plasmodium vinckei*: selection of a strain exhibiting stable resistance to arteether. *Exp Parasitol* 114, 129-132.
- Ranson, H., Jensen, B., Wang, X., Prapanthadara, L., Hemingway, J., Collins, F.H., 2000. Genetic mapping of two loci affecting DDT resistance in the malaria vector *Anopheles gambiae*. *Insect Mol Biol* 9, 499-507.
- Ratcliff, A., Siswantoro, H., Kenangalem, E., Wuwung, M., Brockman, A., Edstein, M.D., Laihad, F., Ebsworth, E.P., Anstey, N.M., Tjitra, E., Price, R.N., 2007. Therapeutic response of multidrug-resistant *Plasmodium falciparum* and *P. vivax* to chloroquine and sulfadoxine-pyrimethamine in southern Papua, Indonesia. *Trans R Soc Trop Med Hyg* 101, 351-359.
- Read, A.F., Lynch, P.A., Thomas, M.B., 2009. How to make evolution-proof insecticides for malaria control. *PLoS Biol* 7, e1000058.
- Read, A.F., Thomas, M.B., 2009. Microbiology. Mosquitoes cut short. *Science* 323, 51-52.
- Reece, S.E., Duncan, A.B., West, S.A., Read, A.F., 2005. Host cell preference and variable transmission strategies in malaria parasites. *Proc Biol Sci* 272, 511-517.

- Reece, S.E., Drew, D.R., Gardner, A., 2008. Sex ratio adjustment and kin discrimination in malaria parasites. *Nature* 453, 609-614.
- Reed, M.B., Saliba, K.J., Caruana, S.R., Kirk, K., Cowman, A.F., 2000. Pgh1 modulates sensitivity and resistance to multiple antimalarials in *Plasmodium falciparum*. *Nature* 403, 906-909.
- Reyburn, H., Mbatia, R., Drakeley, C., Carneiro, I., Mwakasungula, E., Mwerinde, O., Saganda, K., Shao, J., Kitua, A., Olomi, R., Greenwood, B.M., Whitty, C.J., 2004. Overdiagnosis of malaria in patients with severe febrile illness in Tanzania: a prospective study. *Bmj* 329, 1212.
- Roepe, P.D., 2009. Molecular and physiologic basis of quinoline drug resistance in *Plasmodium falciparum* malaria. *Future Microbiol* 4, 441-455.
- Rohrbach, P., Sanchez, C.P., Hayton, K., Friedrich, O., Patel, J., Sidhu, A.B., Ferdig, M.T., Fidock, D.A., Lanzer, M., 2006. Genetic linkage of *pfmdr1* with food vacuolar solute import in *Plasmodium falciparum*. *Embo J* 25, 3000-3011.
- Rosario, V.E., 1976. Genetics of chloroquine resistance in malaria parasites. *Nature* 261, 585-586.
- Rosario, V.E., Hall, R., Walliker, D., Beale, G.H., 1978. Persistence of drug-resistant malaria parasites. *Lancet* 1, 185-187.
- Rosenthal, P.J., Meshnick, S.R., 1996. Hemoglobin catabolism and iron utilization by malaria parasites. *Mol Biochem Parasitol* 83, 131-139.
- Ruebush, T.K., 2nd, Zegarra, J., Cairo, J., Andersen, E.M., Green, M., Pillai, D.R., Marquino, W., Huilca, M., Arevalo, E., Garcia, C., Solary, L., Kain, K.C., 2003. Chloroquine-resistant *Plasmodium vivax* malaria in Peru. *Am J Trop Med Hyg* 69, 548-552.
- Rungsihirunrat, K., Chaijareonkul, W., Seugorn, A., Na-Bangchang, K., Thaithong, S., 2009. Association between chloroquine resistance phenotypes and point mutations in *pfert* and *pfmdr1* in *Plasmodium falciparum* isolates from Thailand. *Acta Trop* 109, 37-40.
- Saito-Nakano, Y., Tanabe, K., Kamei, K., Iwagami, M., Komaki-Yasuda, K., Kawazu, S., Kano, S., Ohmae, H., Endo, T., 2008. Genetic evidence for *Plasmodium falciparum* resistance to chloroquine and pyrimethamine in Indochina and the Western Pacific between 1984 and 1998. *Am J Trop Med Hyg* 79, 613-619.
- Saliba, K.J., Folb, P.I., Smith, P.J., 1998. Role for the *plasmodium falciparum* digestive vacuole in chloroquine resistance. *Biochem Pharmacol* 56, 313-320.
- Sanchez, C.P., Wunsch, S., Lanzer, M., 1997. Identification of a chloroquine importer in *Plasmodium falciparum*. Differences in import kinetics are genetically linked with the chloroquine-resistant phenotype. *J Biol Chem* 272, 2652-2658.
- Sanchez, C.P., McLean, J.E., Stein, W., Lanzer, M., 2004. Evidence for a substrate specific and inhibitable drug efflux system in chloroquine resistant *Plasmodium falciparum* strains. *Biochemistry* 43, 16365-16373.
- Sanchez, C.P., McLean, J.E., Rohrbach, P., Fidock, D.A., Stein, W.D., Lanzer, M., 2005. Evidence for a *pfert*-associated chloroquine efflux system in the human malarial parasite *Plasmodium falciparum*. *Biochemistry* 44, 9862-9870.
- Sanchez, C.P., Rohrbach, P., McLean, J.E., Fidock, D.A., Stein, W.D., Lanzer, M., 2007a. Differences in trans-stimulated chloroquine efflux kinetics are linked to PfCRT in *Plasmodium falciparum*. *Mol Microbiol* 64, 407-420.

- Sanchez, C.P., Stein, W.D., Lanzer, M., 2007b. Is PfCRT a channel or a carrier? Two competing models explaining chloroquine resistance in *Plasmodium falciparum*. *Trends Parasitol* 23, 332-339.
- Sarkari, F., Sheng, Y., Frappier, L., 2010. USP7/HAUSP promotes the sequence-specific DNA binding activity of p53. *PLoS One* 5, e13040.
- Schneider, P., Chan, B.H., Reece, S.E., Read, A.F., 2008. Does the drug sensitivity of malaria parasites depend on their virulence? *Malar J* 7, 257.
- Sharling, L., Sowa, K.M., Thompson, J., Kyriacou, H.M., Arnot, D.E., 2007. Rapid and specific biotin labelling of the erythrocyte surface antigens of both cultured and ex-vivo *Plasmodium* parasites. *Malar J* 6, 66.
- Sharrock, W.W., Suwanarusk, R., Lek-Uthai, U., Edstein, M.D., Kosaisavee, V., Travers, T., Jaidee, A., Sriprawat, K., Price, R.N., Nosten, F., Russell, B., 2008. *Plasmodium vivax* trophozoites insensitive to chloroquine. *Malar J* 7, 94.
- Sidhu, A.B., Verdier-Pinard, D., Fidock, D.A., 2002. Chloroquine resistance in *Plasmodium falciparum* malaria parasites conferred by *pfert* mutations. *Science* 298, 210-213.
- Sidhu, A.B., Valderramos, S.G., Fidock, D.A., 2005. *pfmdr1* mutations contribute to quinine resistance and enhance mefloquine and artemisinin sensitivity in *Plasmodium falciparum*. *Mol Microbiol* 57, 913-926.
- Sidhu, A.B., Uhlemann, A.C., Valderramos, S.G., Valderramos, J.C., Krishna, S., Fidock, D.A., 2006. Decreasing *pfmdr1* copy number in *plasmodium falciparum* malaria heightens susceptibility to mefloquine, lumefantrine, halofantrine, quinine, and artemisinin. *J Infect Dis* 194, 528-535.
- Silva, J.C., Egan, A., Friedman, R., Munro, J.B., Carlton, J.M., Hughes, A.L., 2010. Genome sequences reveal divergence times of malaria parasite lineages. *Parasitology*, 1-13.
- Sina, B.J., Aultman, K., 2001. Resisting resistance. *Trends Parasitol* 17, 305-306.
- Sisowath, C., Petersen, I., Veiga, M.I., Martensson, A., Premji, Z., Bjorkman, A., Fidock, D.A., Gil, J.P., 2009. In vivo selection of *Plasmodium falciparum* parasites carrying the chloroquine-susceptible *pfert* K76 allele after treatment with artemether-lumefantrine in Africa. *J Infect Dis* 199, 750-757.
- Skarbinski, J., Ouma, P.O., Causer, L.M., Kariuki, S.K., Barnwell, J.W., Alaii, J.A., de Oliveira, A.M., Zurovac, D., Larson, B.A., Snow, R.W., Rowe, A.K., Laserson, K.F., Akhwale, W.S., Slutsker, L., Hamel, M.J., 2009. Effect of malaria rapid diagnostic tests on the management of uncomplicated malaria with artemether-lumefantrine in Kenya: a cluster randomized trial. *Am J Trop Med Hyg* 80, 919-926.
- Slater, A.F., Swiggard, W.J., Orton, B.R., Flitter, W.D., Goldberg, D.E., Cerami, A., Henderson, G.B., 1991. An iron-carboxylate bond links the heme units of malaria pigment. *Proc Natl Acad Sci U S A* 88, 325-329.
- Slater, A.F., Cerami, A., 1992. Inhibition by chloroquine of a novel haem polymerase enzyme activity in malaria trophozoites. *Nature* 355, 167-169.
- Soto, J., Toledo, J., Gutierrez, P., Luzz, M., Llinas, N., Cedeno, N., Dunne, M., Berman, J., 2001. *Plasmodium vivax* clinically resistant to chloroquine in Colombia. *Am J Trop Med Hyg* 65, 90-93.
- Srivastava, H.C., Yadav, R.S., Joshi, H., Valecha, N., Mallick, P.K., Prajapati, S.K., Dash, A.P., 2008. Therapeutic responses of *Plasmodium vivax* and P.

- falciparum to chloroquine, in an area of western India where *P. vivax* predominates. *Ann Trop Med Parasitol* 102, 471-480.
- Srivastava, I.K., Morrissey, J.M., Darrouzet, E., Daldal, F., Vaidya, A.B., 1999. Resistance mutations reveal the atovaquone-binding domain of cytochrome b in malaria parasites. *Mol Microbiol* 33, 704-711.
- Srivastava, I.K., Vaidya, A.B., 1999. A mechanism for the synergistic antimalarial action of atovaquone and proguanil. *Antimicrob Agents Chemother* 43, 1334-1339.
- Stepniewska, K., White, N.J., 2008. Pharmacokinetic determinants of the window of selection for antimalarial drug resistance. *Antimicrob Agents Chemother* 52, 1589-1596.
- Su, X., Kirkman, L.A., Fujioka, H., Wellems, T.E., 1997. Complex polymorphisms in an approximately 330 kDa protein are linked to chloroquine-resistant *P. falciparum* in Southeast Asia and Africa. *Cell* 91, 593-603.
- Sumawinata, I.W., Bernadeta, Leksana, B., Sutarnihardja, A., Purnomo, Subianto, B., Sekartuti, Fryauff, D.J., Baird, J.K., 2003. Very high risk of therapeutic failure with chloroquine for uncomplicated *Plasmodium falciparum* and *P. vivax* malaria in Indonesian Papua. *Am J Trop Med Hyg* 68, 416-420.
- Sutherland, C.J., Tanomsing, N., Nolder, D., Oguike, M., Jennison, C., Pukrittayakamee, S., Dolecek, C., Hien, T.T., do Rosario, V.E., Arez, A.P., Pinto, J., Michon, P., Escalante, A.A., Nosten, F., Burke, M., Lee, R., Blaze, M., Otto, T.D., Barnwell, J.W., Pain, A., Williams, J., White, N.J., Day, N.P., Snounou, G., Lockhart, P.J., Chiodini, P.L., Imwong, M., Polley, S.D., 2010. Two nonrecombining sympatric forms of the human malaria parasite *Plasmodium ovale* occur globally. *J Infect Dis* 201, 1544-1550.
- Suwanarusk, R., Russell, B., Chavchich, M., Chalfein, F., Kenangalem, E., Kosaisavee, V., Prasetyorini, B., Piera, K.A., Barends, M., Brockman, A., Lek-Uthai, U., Anstey, N.M., Tjitra, E., Nosten, F., Cheng, Q., Price, R.N., 2007. Chloroquine resistant *Plasmodium vivax*: in vitro characterisation and association with molecular polymorphisms. *PLoS One* 2, e1089.
- Syafruddin, D., Siregar, J.E., Marzuki, S., 1999. Mutations in the cytochrome b gene of *Plasmodium berghei* conferring resistance to atovaquone. *Mol Biochem Parasitol* 104, 185-194.
- Talman, A.M., Domarle, O., McKenzie, F.E., Arie, F., Robert, V., 2004. Gametocytogenesis: the puberty of *Plasmodium falciparum*. *Malar J* 3, 24.
- Tanabe, K., Kato, M., Izumo, A., Hagiwara, A., Doi, S., 1990. *Plasmodium chabaudi*: in vivo effects of Ca²⁺ antagonists on chloroquine-resistant and chloroquine-sensitive parasites. *Exp Parasitol* 70, 419-426.
- Temu, E.A., Kimani, I., Tuno, N., Kawada, H., Minjas, J.N., Takagi, M., 2006. Monitoring chloroquine resistance using *Plasmodium falciparum* parasites isolated from wild mosquitoes in Tanzania. *Am J Trop Med Hyg* 75, 1182-1187.
- Teuscher, F., Gatton, M.L., Chen, N., Peters, J., Kyle, D.E., Cheng, Q., 2010. Artemisinin-induced dormancy in *Plasmodium falciparum*: duration, recovery rates, and implications in treatment failure. *J Infect Dis* 202, 1362-1368.
- Thaithong, S., Suebsaeng, L., Rooney, W., Beale, G.H., 1988. Evidence of increased chloroquine sensitivity in Thai isolates of *Plasmodium falciparum*. *Trans R Soc Trop Med Hyg* 82, 37-38.

- Trampuz, A., Jereb, M., Muzlovic, I., Prabhu, R.M., 2003. Clinical review: Severe malaria. *Crit Care* 7, 315-323.
- Tran, C.V., Saier, M.H., Jr., 2004. The principal chloroquine resistance protein of *Plasmodium falciparum* is a member of the drug/metabolite transporter superfamily. *Microbiology* 150, 1-3.
- Uhlemann, A.C., Cameron, A., Eckstein-Ludwig, U., Fischbarg, J., Iserovich, P., Zuniga, F.A., East, M., Lee, A., Brady, L., Haynes, R.K., Krishna, S., 2005. A single amino acid residue can determine the sensitivity of SERCAs to artemisinins. *Nat Struct Mol Biol* 12, 628-629.
- Valderramos, S.G., Fidock, D.A., 2006. Transporters involved in resistance to antimalarial drugs. *Trends Pharmacol Sci* 27, 594-601.
- Valderramos, S.G., Scanfeld, D., Uhlemann, A.C., Fidock, D.A., Krishna, S., 2010. Investigations into the Role of the *Plasmodium falciparum* SERCA (PfATP6) L263E Mutation in Artemisinin Action and Resistance. *Antimicrob Agents Chemother*.
- van Dijk, M.R., McConkey, G.A., Vinkenoog, R., Waters, A.P., Janse, C.J., 1994. Mechanisms of pyrimethamine resistance in two different strains of *Plasmodium berghei*. *Mol Biochem Parasitol* 68, 167-171.
- Vanderberg, J.P., 2009. Reflections on early malaria vaccine studies, the first successful human malaria vaccination, and beyond. *Vaccine* 27, 2-9.
- Volkman, S.K., Sabeti, P.C., DeCaprio, D., Neafsey, D.E., Schaffner, S.F., Milner, D.A., Jr., Daily, J.P., Sarr, O., Ndiaye, D., Ndir, O., Mboup, S., Duraisingh, M.T., Lukens, A., Derr, A., Stange-Thomann, N., Waggoner, S., Onofrio, R., Ziaugra, L., Mauceli, E., Gnerre, S., Jaffe, D.B., Zainoun, J., Wiegand, R.C., Birren, B.W., Hartl, D.L., Galagan, J.E., Lander, E.S., Wirth, D.F., 2007. A genome-wide map of diversity in *Plasmodium falciparum*. *Nat Genet* 39, 113-119.
- von Seidlein, L., Duraisingh, M.T., Drakeley, C.J., Bailey, R., Greenwood, B.M., Pinder, M., 1997. Polymorphism of the *Pfmdr1* gene and chloroquine resistance in *Plasmodium falciparum* in The Gambia. *Trans R Soc Trop Med Hyg* 91, 450-453.
- Walliker, D., Carter, R., Sanderson, A., 1975. Genetic studies on *Plasmodium chabaudi*: recombination between enzyme markers. *Parasitology* 70, 19-24.
- Walliker, D., Quakyi, I.A., Wellems, T.E., McCutchan, T.F., Szarfman, A., London, W.T., Corcoran, L.M., Burkot, T.R., Carter, R., 1987. Genetic analysis of the human malaria parasite *Plasmodium falciparum*. *Science* 236, 1661-1666.
- Walliker, D., Hunt, P., Babiker, H., 2005. Fitness of drug-resistant malaria parasites. *Acta Trop* 94, 251-259.
- Wang, J., Huang, L., Li, J., Fan, Q., Long, Y., Li, Y., Zhou, B., 2010. Artemisinin directly targets malarial mitochondria through its specific mitochondrial activation. *PLoS One* 5, e9582.
- Warhurst, D.C., 2003. Polymorphism in the *Plasmodium falciparum* chloroquine-resistance transporter protein links verapamil enhancement of chloroquine sensitivity with the clinical efficacy of amodiaquine. *Malar J* 2, 31.
- Wellems, T.E., Panton, L.J., Gluzman, I.Y., do Rosario, V.E., Gwadz, R.W., Walker-Jonah, A., Krogstad, D.J., 1990. Chloroquine resistance not linked to *mdr*-like genes in a *Plasmodium falciparum* cross. *Nature* 345, 253-255.

- Wellems, T.E., Walker-Jonah, A., Panton, L.J., 1991. Genetic mapping of the chloroquine-resistance locus on *Plasmodium falciparum* chromosome 7. *Proc Natl Acad Sci U S A* 88, 3382-3386.
- Wenger, J.W., Schwartz, K., Sherlock, G., 2010. Bulk segregant analysis by high-throughput sequencing reveals a novel xylose utilization gene from *Saccharomyces cerevisiae*. *PLoS Genet* 6, e1000942.
- White, N.J., Olliaro, P.L., 1996. Strategies for the prevention of antimalarial drug resistance: rationale for combination chemotherapy for malaria. *Parasitol Today* 12, 399-401.
- White, N.J., 2002. The assessment of antimalarial drug efficacy. *Trends Parasitol* 18, 458-464.
- WHO, W.H.O., 2006. Guidelines for the treatment of malaria, second edition.
- WHO, W.H.O., 2008. World Malaria Report 2008.
- WHO, W.H.O., 2009. Methods for surveillance of antimalarial drug efficacy.
- Witkowski, B., Lelievre, J., Barragan, M.J., Laurent, V., Su, X.Z., Berry, A., Benoit-Vical, F., 2010. Increased tolerance to artemisinin in *Plasmodium falciparum* is mediated by a quiescence mechanism. *Antimicrob Agents Chemother* 54, 1872-1877.
- Wootton, J.C., Feng, X., Ferdig, M.T., Cooper, R.A., Mu, J., Baruch, D.I., Magill, A.J., Su, X.Z., 2002. Genetic diversity and chloroquine selective sweeps in *Plasmodium falciparum*. *Nature* 418, 320-323.
- Yang, Y.Z., Asawamahesakda, W., Meshnick, S.R., 1993. Alkylation of human albumin by the antimalarial artemisinin. *Biochem Pharmacol* 46, 336-339.
- Zhang, H., Howard, E.M., Roepe, P.D., 2002. Analysis of the antimalarial drug resistance protein PfCRT expressed in yeast. *J Biol Chem* 277, 49767-49775.
- Zhang, H., Paguio, M., Roepe, P.D., 2004. The antimalarial drug resistance protein *Plasmodium falciparum* chloroquine resistance transporter binds chloroquine. *Biochemistry* 43, 8290-8296.
- Zhou, Z., Griffing, S.M., de Oliveira, A.M., McCollum, A.M., Quezada, W.M., Arrospide, N., Escalante, A.A., Udhayakumar, V., 2008. Decline in sulfadoxine-pyrimethamine-resistant alleles after change in drug policy in the Amazon region of Peru. *Antimicrob Agents Chemother* 52, 739-741.